PB83-127738

Sampling Frequency
Microbiological Drinking Water Regulations

Drexel Univ.
Philadelphia, PA

Prepared for

Environmental Protection Agency
Washington, DC

Sep 82

U.S. Department of Commerce
National Technical Information Service
NTIS

United States
Environmental Protection
Agency

Office of
Drinking Water (WH-550)
Washington DC 20460

September 1982
EPA/570/9-82-001

PB83-127738

EPA

Sampling Frequency-Microbiological Drinking Water Regulations

Final Report

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

BIBLIOGRAPHIC INFORMATION

PB83-127738

Sampling Frequency - Microbiological Drinking Water Regulations.

Sep 82,

Wesley O. Pipes, and Robert R. Christian.

PERFORMER: Drexel Univ., Philadelphia, PA.
Grant EPA-R-805637

SPONSOR: Environmental Protection Agency, Washington, DC.
Office of Drinking Water.
EPA/570/982-001
Final rept.

The primary goal of this project was to develop a sampling model which can be used to specify the sampling frequency needed to determine compliance with the microbiological maximum contaminant levels of the National Interim Primary Drinking Water Regulations. Two approaches to model development were used. The emperical approach was based on fitting coliform data to frequency distributions. The mechanistic approach was an attempt to analyze the physical elements of the water system in order to determine if samples from a particular location are representative of water quality over an area and if it is more likely that coliforms will be found in certain types of sampling locations.

KEYWORDS: *Potable water, *Bacteria, *Drinking water, *Water pollution detection.

Available from the National Technical Information Service, Springfield, Va. 22161

PRICE CODE: PC A09/MF A01

SAMPLING FREQUENCY - MICROBIOLOGICAL
DRINKING WATER REGULATIONS
Final Report

by

Wesley O. Pipes and Robert R. Christian
Department of Biological Sciences
Drexel University
Philadelphia, Pennsylvania

EPA R-805-637

Project Officers

Paul Berger	Edwin E. Geldreich
Criteria & Standards Division	Water Supply Research Division
Office of Drinking Water	Municipal Environmental Research Lab.
Washington, D.C. 20460	Cincinnati, Ohio 45268

OFFICE OF DRINKING WATER
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D. C. 20460

DISCLAIMER

This report has been reviewed by the Office of Drinking Water and the Municipal Environmental Research Laboratory, U.S. Environmental Protection Agency and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names of commercial products constitute endorsement or recommendation for use.

ABSTRACT

The primary goal of this project was to develop a sampling model which can be used to specify the sampling frequency needed to determine compliance with the microbiological maximum contaminant levels of the National Interim Primary Drinking Water Regulations. Two approaches to model development were used. The emperical approach was based on fitting coliform data to frequency distributions. The mechanistic approach was an attempt to analyze the physical elements of the water system in order to determine if samples from a particular location are representative of water quality over an area and if it is more likely that coliforms will be found in certain types of sampling locations.

The emperical approach resulted in the finding that coliform counts in samples from a water distribution system do not fit either Poisson or Poisson plus added zeros distributions. The counts do fit negative binomial distributions and truncated lognormal distributions. It was proposed to use the truncated lognormal distribution because it is familiar to personnel in the water works field and computations are somewhat simpler. Truncation in this case means that coliform densities <1/100ml or >80/100ml cannot be measured because of limitations of the laboratory methodology for coliform detection and enumeration. The probability of violation of one of the microbiological maximum contaminant levels is a function not only of the densities of coliforms in the water and the number of samples collected but also of the parameters of the lognormal distribution. Higher values of the geometric standard deviation (that is, a higher degree of aggregation of coliform bacteria in the system) lead to larger numbers of samples being needed for detection of contamination.

The mechanistic approach resulted in the finding that water distribution systems can be divided into hydraulically isolated sections. Coliform occurrence can vary significantly among the isolated sections of a water distribution system which leads to the conclusion that all sections of the system must be included in the microbiological monitoring program in order to get an accurate measure of coliform occurrence. No significant difference in the occurrence of coliforms between peripheral and nonperipheral sampling locations was found. Also, no increase in coliform occurrence with distance from the water source into the system was found. Thus, sampling locations for a microbiological monitoring program should be selected by a randomization procedure which gives equal probability of selection to all sampling locations.

CONTENTS

Foreword iii
Abstract iv
Figures viii
Tables x
List of Abbreviations and Symbols xiv
Acknowledgments xvi

1. Introduction 1
    1.1 Objectives of the Project 1
    1.2 Approaches 1
    1.3 Objectives of Monitoring 2
        1.3.1 Safety 2
        1.3.2 Reliability 3
2. Conclusions and Recommendations 6
3. Methods and Procedures 8
    3.1 Introduction 8
    3.2 Planning a Sampling Program 8
        3.2.1 Hierarchy of Sampling Units 8
        3.2.2 Random Sampling 12
        3.2.3 Stratified Random Sampling 15
        3.2.4 Other Sampling Patterns 15
        3.2.5 Application to Monitoring Programs 16
    3.3 Collection and Analysis of Samples 16
        3.3.1 Collection of Samples 17
        3.3.2 Analysis of Samples 17
        3.3.3 Special Quality Assurance Procedures 19
        3.3.4 Data Management 19
4. Sample Validity 22
    4.1 Introduction 22
    4.2 Subsequent Contamination 23
        4.2.1 Sterility of Sample Bottles........ 23
        4.2.2 Tap Location 23
        4.2.3 Sampler Error 27
        4.2.4 Contamination During Transit in Ice 28
        4.2.5 Carry Over During Filtration 30
        4.2.6 Summary on Subsequent Contamination 35
    4.3 Changes in Coliform Densities in Water Samples 36
        4.3.1 Preservation Techniques 36
        4.3.2 Second Day Filtration 37
    4.4 Observations on Interferences 40
    4.5 Verifications of Presumptive Coliform and Other Colonies .. 43

5. Maximum Contaminant Levels for Coliform Bacteria and Their Relation to Frequency Distributions 54
  - 5.1 Introduction 54
  - 5.2 Characteristics of MCL's Independent of Specific Frequency Distributions 54
  - 5.3 Previous Work on Frequency Distributions of Bacteria 56
    - 5.3.1 Distributions Considered 56
    - 5.3.2 Poisson Distribution 56
    - 5.3.3 Poisson Plus Added Zeroes 60
    - 5.3.4 Assumptions Underlying the Simplistic Models 60
    - 5.3.5 Lognormal Distribution 63
    - 5.3.6 Negative Binomial Distribution 67
  - 5.4 Data Fitting 67
    - 5.4.1 Data Sets 70
    - 5.4.2 Poisson Distribution 70
    - 5.4.3 Poisson Plus Added Zeroes Distribution 70
    - 5.4.4 Lognormal Distribution 74
    - 5.4.5 Negative Binomial Distribution 83
  - 5.5 Relationship of the Lognormal Distribution to the National Interim Primary Drinking Water Regulations 85
    - 5.5.1 General 85
    - 5.5.2 Probability of Exceeding the Average Rule 86
    - 5.5.3 Probability of Exceeding the Maximum Rule 91
    - 5.5.4 Probability of Obtaining Samples above Particular Densities 94
    - 5.5.5 The Probability of Exceeding Regulations if the Number of Samples is Reduced 99
    - 5.5.6 Interpreation of Field Results When All Samples are Negative 99
  - 5.6 Replication of Samples and Larger Volume Samples 100
    - 5.6.1 Introduction 100
    - 5.6.2 Replicate Samples Collected 101
    - 5.6.3 Results from Duplicate Samples 101
    - 5.6.4 Negative Sampling Results 111
    - 5.6.5 Positive Sampling Results 112
    - 5.6.6 Larger Volume Samples 112

6. Mechanistic Approach 125
  - 6.1 Introduction 125
  - 6.2 Configuration of Sections 125
    - 6.2.1 Definitions of Distribution System Elements 125
    - 6.2.2 Isolated Sections 126
    - 6.2.3 Configuration of Systems Sampled 128
    - 6.2.4 Relationship of Configuration to Sampling 131
  - 6.3 Sampling by Sections 134
    - 6.3.1 Differences Among Sections 134
    - 6.3.2 Differences Within Sections 143
  - 6.4 Hypotheses About Contamination 145
    - 6.4.1 Upstream and Downstream Samples 150
    - 6.4.2 Peripheral Locations 152
    - 6.4.3 Distance from Treatment Plant 153
  - 6.5 Summary and Implications 156

Appendices (Appendices are not included in this publication, but are available from the Office of Drinking Water.)

A. Coatesville, Pennsylvania
B. Woodbury Heights, New Jersey
C. Long Beach Water Co.
D. Downingtown, Pennsylvania
E. Brooklawn, New Jersey
F. Mt Idy Mobile Home Park
G. Spring Run Homes
H. Marshallton Woods Homes
I. Bradford Glen Homes
J. Effect of Filter Backwashing on Bacteriological Quality of Drinking Water - (included in this publication)

FIGURES

Number		Page
3.1	Woodbury Heights Sampling Location Map	10
3.2	Portion of Location File from Woodbury Heights, N.J.	20
3.3	Portion of Bacteriological File from Woodbury Heights, N.J.	21
5.1	Probability of Exceeding Average Rule for Random Dispersion of Coliforms at Different Actual Mean Densities	58
5.2	Probability of Exceeding Maximum Rule for Random Dispersion of Coliforms in a Distribution System	59
5.3	Probability of Obtaining No Positive Samples for Actual Mean Density of 1 per 100ml	61
5.4	Distribution of computed arithmetic mean (μ) from the geometric means and geometric standard deviations in lognormal distributions. Isobars represent different arithmetic means.	64
5.5	The percentage of samples with $\geq$ coliforms/100ml as a function of the geometric mean (Gm) and geometric standard deviation.	65
5.6	Likelihood of exceeding the average and maximum rules for lognormal distributions when a large number of samples are taken per month.	66
5.7	The Probability of Obtaining a Sample Mean of 1 for a True Mean Density of 1 as a Function of Number (n) of Samples and Variability of the System. (after Muenz, 1978).	68
5.8	Probability of obtaining a particular count (X) when $\mu=1$ and K is varied from 0.1 to 10 in a negative binomial distribution. X=0, 1, 2, 3 $\geq$ 4.	69
5.9	Log-Probability Plots of Coatesville Coliform Data	75
5.10	Extrapolation of Log-Probability Plot of Coatesville Data for Fitting a Truncated Lognormal Distribution.	76
5.11	Probability of Exceeding the Average Rule when $\mu=1$ as a Function of Number of Samples	88
5.12	Probability of Exceeding the Average Rule when $\alpha=1$ as a Function of Number of Samples	90
5.13	Probability of Exceeding the Maximum Rule when $\alpha=1$ as a Function of Number of Samples	93

Number Page

5.14 Probability of Obtaining at Least 1 Positive Sample as a Function of Number of Samples when $\alpha = 1$. 96

5.15 Probability of Obtaining at Least 1 Sample with $\geq$ 4 Coliforms per 100ml when $\alpha = 1$ as a Function of Number of Samples 97

5.16 Probability of Obtaining at Least One TNTC Sample when $\alpha = 1$ as a Function of Number of Samples 98

6.1 Representative Examples of Four Isolated Section Configurations 127

6.2 Relationship of average cross sectional area to volume of water distribution system sections. 132

6.3 Flow Diagrams for the Isolated Sections of System BL. 151

TABLES

Number		Page
3.1	Sections of Woodbury Heights, N.J.	11
4.1	Comparison of Positive Coliform Results in Source Water and Distribution System Samples for Seven Systems	24
4.2	Comparison of Coliform Results from Different Sampling Sites	25
4.3	Comparison of Paired Samples from Commercial Establishments	26
4.4	Tests for Subsequent Contamination from Fingers of Samplers	28
4.5	Contamination of Sample Bottles Immersed in Ice Water Containing Different Densities of Coliforms	29
4.6	Positive Samples from Different Types of Bottles Immersed in Ice Water for Various Periods of Time	31
4.7	Relationships of Coliform Density, Type of Filter Funnel and Rinsing Protocol as Determinants of Carry Over	33
4.8	Potential Carry Over Analysis of Selected Samples Collected in the Field	34
4.9	Effect of Preservation Techniques on Standard Plate Count	38
4.10	Comparison of Preservation Techniques	39
4.11	Delayed Filtration Experiments	41
4.12	Verification Results on Systems Sampled	45
4.13	Summary of Verifications and API Identifications	46
4.14	Summary of Reactions of Isolates Identified as Enterobacter Agglomerans (IND+)	49
4.15	Summary of Reactions of Isolates Identified as Enterobacter Agglomerans (IND-)	50
4.16	Summary of Isolates Identified as E. Coli	51
4.17	Comparison of Test Results Between Escherichia Coli and Enterobacter Agglomeran Strains which are IMViC Type ++--	52
4.18	Summary of API Identification of Non-Coliform Colonies	53
5.1	Examples of Various Violations of the MCL's when Sample Size is 10	55
5.2	Summary of Water Systems Sampled	71

Number Page

5.3 Applicability of the Poisson Distribution to the Frequency Distribution of Confirmed Total Coliforms in Water Distribution Systems 72

5.4 Applicability of the Poisson Plus Added Zeroes Distribution to the Frequency Distribution of Confirmed Total Coliforms in Water Distribution Systems 73

5.5 Example of Computation of Factors to be Used in the Assessment of the Lognormal Distribution Based on Data from WH I 78

5.6 Parameters of the Lognormal Distribution of Coliforms in Woodbury Heights, N.J. as Computed by the Graphical Approach using all Locations Having Equivalent Volumes Filtered 79

5.7 Parameters of the Lognormal Distribution of Coliforms in Water Distribution Systems as Computed by Least Squares Regression 80

5.8 Relationship of Sample Arithmetic Mean ($\bar{x}$) and Computed Arithmetic Means from the Regression Analysis (αr) and Graphical Analysis (αg) for the Lognormal Distribution

5.9 Parameters of the Negative Binomial Distribution of Confirmed Total Coliforms in Water Distribution Systems 84

5.10 Probability of Exceeding the Average Rule as a Function of Number of Samples when the Geometric Mean is Held Constant 86

5.11 Computed Arithmetic Means (α) When $\mu = 1$ as a Function of Number of Samples 87

5.12 The Average Computed Arithmetic Mean (α) When $\mu = 0$ as a Function of Number of Samples 89

5.13 Probability of Exceeding the Maximum Rule as a Function of Number of Samples When Geometric Mean is Held Constant 92

5.14 Occurrence of Violations in the Average and Maximum Rules in Water Distribution Systems Studied 94

5.15 Sequences of Coliform Counts in Replicate Samples-Coatesville 102

5.16 Sequences of Coliform Counts in Replicate Samples-Woodbury Heights 103

5.17 Sequences of Coliform Counts in Replicate Samples-Long Beach Water Co. 104

5.18 Sequences of Coliform Counts in Replicate Samples-Downingtown 105

5.19 Sequences of Coliform Counts in Replicate Samples-Brooklawn 106

5.20 Sequences of Coliform Counts in Replicate Samples-Mt. Idy Mobile Home Park 107

5.21 Sequences of Coliform Counts in Replicate Samples-Spring Run 108

5.22 Sequences of Coliform Counts in Replicate Samples-Marshallton Woods 109

Number		Page
5.23	Sequences of Coliform Counts in Replicate Samples-Bradford Glen	110
5.24	Analysis of Data from Duplicate 100ml Samples	113
5.25	Effect of Second Replicate 100ml Sample on Fraction of Sites Positive and Percent of Sites with 4 or more Coliforms per 100ml	114
5.26	Comparison of 100ml and 200ml Samples	115
5.27	Expected Results of Sampling from Lognormal Distributions	118
5.28	Expected Numbers of Counts in Samples from Middle Range Coliform Density Sites	119
5.29	Expected Results of Sampling from Lognormal Distributions	120
5.30	Effect of Sample Volume on Fraction of Sites with Coliforms	122
5.31	Effect of Sample Volume on Fraction of Sites with Coliforms	123
6.1	Structural Analysis of Systems Sampled	129
6.2	Summary of Structural Parameters by Section Configuration	130
6.3	Dimensions of Distribution System Pipes by Section	133
6.4	Analysis of System CV Bacteriological Water Quality Parameters by Section	135
6.5	Analysis of System WH Bacteriological Water Quality Parameters by Section	136
6.6	Analysis of System LB Bacteriological Water Quality Parameters by Section	137
6.7	Analysis of System DT Bacteriological Water Quality Parameters by Section	138
6.8	Analysis of System BL Bacteriological Water Quality Parameters by Section	139
6.9	Analysis of System SR Bacteriological Water Quality Parameters by Section	139
6.10	Contingency Table Analyses for Differences among Sections	141
6.11	Analysis of System LB Data for June 1979	142
6.12	Analysis of System WH Data for May-June 1981	144
6.13	Comparison of Sampling Results (100ml Samples) from the Two Locations of System MW	146
6.14	Comparison of Sampling Results (100ml Samples) from 68 Sampling Locations of System BL	147
6.15	Comparison of Sampling Results (100ml Samples) from 18 Sampling Locations of System SR	149

Number		Page
6.16	Locations of Positive Coliform Samples Collected from System BL - June 1981	152
6.17	Peripheral Versus Other Locations	153
6.18	Analysis of Coliform Occurrence as a Function of Distance from the Treatment Plant System BL - June 1981 Data	154
6.19	Analysis of Coliform Occurrence as a Function of Distance from the Treatment Plant System WH - May-June 1981	155
6.20	Analysis of Coliform Occurrence as a Function of Distance from the Treatment Plant System LB	155

LIST OF ABBREVIATIONS AND SYMBOLS

ABBREVIATIONS

BB -- Brant Beach service area of the Long Beach Water Co. distribution system
BG -- Bradford Glen water system
BL -- Brooklawn water system
CFU -- colony forming units
CV -- Coatesville water system
DT -- Downingtown water system
gpd -- gallons per day
LB -- Long Beach Water Co. water system
MCL -- maximum contaminant level
MGD -- million gallons per day
MI -- Mount Idy water system
MW -- Marshallton Woods water system
SR -- Spring Run water system
TNTC -- too numerous to count
TP -- Terrace Plant service area of the Long Beach Water Co. distribution system
WH -- Woodbury Heights water system

SYMBOLS

D_{max}	Statistic used for Kolmogorov-Smirnov test
D^2	Fisher's index of dispersion
e	Base of natural logarithms
F	Statistic used for regression analysis
GM	Geometric mean
GSD	Geometric standard deviation
GSE	Geometric standard error
k	Parameter of negative binomial distribution
n	Number of samples
P	Probability
R^2	Coefficient of Determination
S	Standard deviation
S^2	Variance
S_g	Geometric standard deviation
X	Coliform count of a single sample
$\bar{x}$	Mean coliform count of a set of samples
$\bar{X}_g$	Geometric mean coliform count
Z	Statistic used for normal distribution

SYMBOLS (con't)

α	Arithmetic mean calculated from parameters of lognormal distribution
α_g	Arithmetic mean calculated from parameters of lognormal distribution estimated by graphical technique
α_r	Arithmetic mean calculated from parameters of lognormal distribution estimated by regression
ζ_2	Fraction of water not contaminated with coliform bacteria
χ^2	Statistic used for goodness-of-fit and contingency table analyses
σ	Variance
μ	Mean

ACKNOWLEDGMENTS

The basic questions to which this project was directed were formulated by Joseph A. Cotruvo and Charles W. Hendricks, Office of Drinking Water, U.S. Environmental Protection Agency. Dr. Hendricks provided guidance as Project Officer during the first year of the project and maintained his scientific interest in the research after his move to the Office of Research and Development. Paul Berger, Office of Drinking Water and E. E. Geldreich, Municipal Environmental Research Laboratory served as Project Officers during the last two years of the project and made significant intellectual contributions.

Several individuals in the Pennsylvania Department of Environmental Resources and the New Jersey Department of Environmental Protection provided information which was essential for the identification and selection of systems for sampling. We would also like to express appreciation to the City of Coatesville, Pennsylvania, Woodbury Heights Township, New Jersey, The Long Beach Water Company, Downingtown Borough, and Brooklawn Township, New Jersey for their cooperation in the sampling programs.

Harvey A. Minnigh supervised the sampling programs and laboratory analyses for the last two years of the project. Edward M. Podgorski was responsible for the laboratory studies during the first year of the project. Marc Goshko, Eric D. Becker, James Sioma, Gary Burlingame, Marlene Troy and others participated in the data collection and analysis. Mrs. Sandra L. Sees typed and retyped numerous reports and manuscripts and was an unfailing source of cheerful encouragement.

SECTION 1

INTRODUCTION

1.1 OBJECTIVES OF THE PROJECT

The overall objective of the project was to obtain data and information for use in developing a scientifically verifiable rationale for specifying monitoring programs for water systems of different sizes and types which can be used to determine compliance with the maximum microbiological contaminant level (MCL) of the National Interim Primary Drinking Water Regulations. Among the items specified in the Regulations are 1) the sampling frequency in terms of number of samples per month as a function of population served, 2) the standard sample, and 3) a statement that the samples shall be taken at points which are representative of conditions in the distribution system.

The primary goal of the project was to develop a sampling model which could be used for revision of the sampling frequency table. The data base for the sampling model is the results from short term sampling of nine small, community water systems. The collection of field sampling data for the data base was limited by physical considerations to small communities within about 75 miles of Philadelphia. There may be special water quality conditions which occur in this region and which are not of any great significance in other regions or there may be phenomena occurring in other regions which are not pertinent here. Thus, the sampling model needs to be verified for other regions of the country.

The data for model development were collected using the membrane filter method for determining coliform densities of water samples. No information on the multiple tube technique are presented in this report.

1.2 APPROACHES

Two approaches to the construction of a sampling model were used. One, which we call the empirical approach, is based on fitting various data sets to frequency distributions. The frequency distributions are then used to predict the probabilities of obtaining various results with smaller numbers of samples. This approach is explored in Section 5.

The other, which is called the mechanistic approach, depends upon postulating mechanisms by which the coliform bacteria are despersed in the distribution system. If the mechanism of contamination and the details of flow and mixing in the distribution system are known, it should be possible to predict where the coliforms would be found and possibly the frequency distribution for the coliform counts. This approach is explored in Section 6.

The use of either approach to the construction of a sampling model requires accurate sampling data for calibration. It is necessary to avoid erroneous results; that is, coliform bacteria which were not actually in the distribution system showing up in the laboratory tests or coliform bacteria in the distribution system not showing up as coliform bacteria in the laboratory analyses. The special procedures which were used for sample collection and handling are described in Section 3 and questions about sample validity are addressed in Section 4.

1.3 OBJECTIVES OF MONITORING

The overall objective of microbiological monitoring of water distribution systems is related to the protection of public health, especially to the prevention of the spread of waterborne diseases. Historically, particularly in the nineteenth century, there were many major epidemics of diseases transmitted by water contaminated by human wastes. Prevention of waterborne disease depends upon multiple barriers; i.e., selection of the waters of the highest quality available for public supply, protection of the water source, treatment to remove turbidity particles, disinfection to kill pathogens, maintaining a residual concentration of disinfectant in the distribution system, and protection of the physical integrity of the distribution system against cross connections and leaks into the system.

The most difficult barrier to maintain, and the one whose penetration poses the greatest threat of waterborne disease, is the physical integrity of the distribution system. All water distribution systems leak and potential cross connections are not uncommon. Any water which leaks into the system may contain pathogenic microorganisms. Positive pressure in the distribution system should assure that the leakage is always out of the system but pressure variations due to differences in elevation, very high rates of water use for fighting fires, excessively high demand during hot weather, and accidental breaks can result in significantly reduced water pressure and enhance the risk of contaminated water leaking into the system.

Outbreaks of waterborne disease are infrequent in any one system but twenty to thirty such outbreaks are reported in the United States every year. Thus, it appears that the multiple barrier approach as presently practiced is effective but not absolute. Microbiological monitoring of the product delivered to the consumer is one method of directing attention to the necessity of constant attention and maintenance of the barriers.

Harrison (1978) made the distinction between the safety of an individual portion of the water and the reliability of the water system. We contend that the routine microbiological monitoring of water distribution systems as required by the National Interim Primary Drinking Water Regulations is directed to demonstrating the reliability of the overall system and that the safety of individual portions of the water is not an issue.

1.3.1 Safety.

Although the total coliform group is not an infallible indicator, the demonstration of the absence of coliform bacteria from some portion of water from a distribution system does provide some reasonable assurance of the safety of that water (in terms of a low probability of transmission of waterborne infectious disease). Unfortunately, it is impractical to test any significant fraction of the water which will be consumed by humans to achieve this reasonable assurance of safety. It is quicker, cheaper, and more effective to boil and cool water before consumption to achieve the assurance of safety than to to monitor for microorganisms and that is exactly what is done for particularly susceptable individuals such as young infants and those with chronic, debilitating diseases.

As an example of the safety issue, consider a city block, 400 feet long with a 6" diameter main under the street. At any time, there are about 588 gallons or 2224 liters of water in the main. Suppose that on a given day samples are collected from two residences at either end of the block and that the samples each have a volume of 250ml. In the laboratory the samples are well shaken and 100ml of each is filtered. If all goes well, there will be a demonstration 22 hours later that no coliforms were present in either sample. (The possibility that 1 or 2 coliforms might have been present in the 250ml of the sample but were not included in the 100ml which were filtered has to be admitted but is of little consequence.) This does not provide proof that there were no coliforms in the other 2223.5 liters of water in the main between the two sampling points and, even if it did, by the time the coliform results are available a different 2224 liters of water is in the main. The water in the samples tested is never consumed and the water which is consumed is never tested; thus, the safety of the water consumed is never demonstrated.

The sampling frequency assumed in this example is extreme. System WH (Appendix B) served a population of 3600 distributed over 115 blocks from which samples could be obtained. Four samples per month were examined for routine monitoring. If the sites to be sampled were selected randomly, each block would be sampled once every 29 months on the average rather than the two samples in one day of the example. Actually, some blocks are never sampled and some blocks are sampled more frequently than once every 29 months. In any case, it is clear that confidence can be placed in the results of one sample from a given block every 2 or 3 years only if the barriers against microbial contamination are known to be reliable.

1.3.2 Reliability.

The primary objective of a microbiological monitoring program is to demonstrate the reliability of multiple barriers against transmission of waterborne disease. This reliability is not absolute. It must be admitted that for any water supply there will always be conditions under which the barriers will be penetrated. The reliability of the system needs to be measured quantitatively in order to determine the extent of protection of public health.

The techniques available for testing for microorganisms require that discrete volumes of water be used. The standard sample for the membrane filter techniques is 100ml at present. Thus, the potential number of samples is all 100ml volumes of water which pass through the distribution system. For a one million gallon per day (MGD) system, this would amount to 37,800,000 potential samples per day or 1.134×10^9 samples per month. The potential number of samples which could be tested is, for all practical purposes, infinite in relation to the actual number of samples tested. The question to be considered is what inferences can be drawn about the reliability of the system from the few samples which are collected and tested.

Reliability is a characteristic which is related to persistence in time and over the area served by the distribution system. A distribution system which is free of coliforms one day but has many the next would not be considered to be reliably protected. Likewise, a system free of coliforms in some areas but with many of them in other areas would not be considered to be reliably protected. Thus, in monitoring to demonstrate reliability of protection the sampling must extend over time and over the entire area covered by the distribution system. The need for the monitoring program to be representative of the entire area of the distribution system is obvious. The question of time persistence of reliability needs additional consideration.

As an example of the reliability question, consider a system supplying about 450,000 gallons per day (gpd) to a population of 4500. Under present regulations, 5 samples per month would be examined for microbiological contamination. Assume that the water system has certain characteristics of water source, treatment, and the nature of the distribution system such that, over some relatively long period of time, some fraction of the 100ml samples will have coliforms present. For any month, if all 5 samples have no coliforms, it can be said that we are 95% confident that the fraction of the 100ml samples with coliforms is less than one half (0.5) for that month (see section 5.3.2). However, if all of the 60 samples collected in a year's time have no coliforms, it can be said that we are 95% confident that the fraction of 100ml samples with coliforms is less than one twentieth (0.05) for that year, and if no coliforms are found in the 300 samples collected over a five year period, it can be said that we are 95% confident that the fraction of 100ml samples with coliforms is less than one hundreth (0.01) over the five year period.

The assumptions used for this example need to be examined carefully. The accumulation of negative coliform results over a period of time developes more and more confidence that the barriers to microbial contamination of the water are reliable. However, it is not clear how long a period of time the characteristics of the system remain the same so that data may be accumulated to demonstrate reliability of protection. The coliform density in a water distribution system can change from one day to the next and almost certainly will change over the seasons of a year. These changes can be considered to be normal variability resulting from the characteristics of the water system. The characteristics of the system may change when there are alterations in the water source, treatment processes, or distribution system. Examples of these alterations would include changes in water treatment processes, placing new water mains in service, failure and replacement of existing water mains, storage tank painting and repairs, and abnormal variations in water pressure. These are identifiable occurrences and can be used to separate groups of data used to demonstrate reliability of the system.

Even in the absence of identifiable changes in the water system, it may be expected that there will be slow deterioration, particularly of the distribution system, which would change the reliability of protection against microbial contamination. In the preceding example, time periods of one month, one year, and five years were used. It seems likely that the characteristics of a water system would remain consistent for more than one month but probably not for as long as five years. It seems reasonable to evaluate the reliability of protection of a water supply against microbiological contamination on the basis of data collected over one or two year periods in the absence of identifiable occurrences which could change the characteristics of the system.

There is no hard evidence to show that the reliability of protection of a water system should be evaluated using data collected over a one year or two year or five year period. In this project, water systems were sampled over one or two month periods and the data collected shed no light on this question. There is, however, a precedent for evaluating the reliability of protection over a one year period under the previous drinking water standards of the U.S. Public Health Service (1962). Water supplies had "approval" if they met the coliform standards all 12 months of a year, "provisional approval" if the standards were violated only once in 12 months, and were "use prohibited" if the standards were violated any two months in a year's time.

Clearly, averaging coliform results over a month's time is a matter of convenience for reporting and has no particular scientific basis. However, there is no evidence that hourly, daily, weekly or quarterly averaging and reporting would give a superior measure of coliform occurrence and densities in the system. Since the monitoring and reporting system is run by people for the protection of people, it is appropriate to select the reporting period as a matter of convenience in the absence of scientific evidence that one reporting period would be superior to any other. The same argument can be made in respect to the period of time over which the reliability of the system should be evaluated.

SECTION 2

CONCLUSIONS AND RECOMMENDATIONS

1. The reliability of any water system in providing properly treated and disinfected water is never absolute. The objective of microbiological monitoring of a water distribution system is to provide a quantitative measure of this reliability which can only be expressed as a probability function.

2. Contamination of a water sample with coliforms during or subsequent to collection can be prevented or recognized by a proper quality assurance program.

3. When coliforms are present in a water sample collected from a distribution system, their density can change within twenty-four hours if the sample is held that long before examination. Thus, there may be a transit time problem using currently accepted procedures.

4. Coliforms were found in all nine community water systems studied. The most common verified species were *Enterobacter cloacae*, *Klebsiella pneumoniae*, and *Klebsiella oxytoca*. *Escherichia coli* was found only rarely and only in two systems.

5. Most of the presumptive coliform colonies observed on membrane filters which did not verify were strains of *Enterobacter agglomerans*. However, a sucrose positive strain of this species which did produce gas in lauryl tryptose and brilliant green bile broths was encountered. This strain has the same IMViC type as *Escherichia coli*.

6. Coliform counts from all systems studied could be fitted to either the truncated lognormal or the negative binomial distribution. This demonstrates that coliforms are *not* randomly dispersed in water distribution systems. The variance of the counts was much greater than the mean.

7. We recommend the truncated lognormal distribution for modeling coliform counts in samples from a distribution system because some water works personnel are already familiar with it and the estimation of parameters is mathematically simpler.

8. Even when the mean coliform density is greater than 1 per 100ml, the probability that a 100ml sample will have no coliforms present is very high. This is a prediction from the truncated lognormal distribution which was very well supported by the data collected.

9. Most of the systems studied had coliform counts with geometric means in the range 10^{-1} to 10^{-4} per 100ml and geometric standard deviations between 10 and 100. The average arithmetic coliform counts in the samples collected from a system ranged from 0.1 to > 9.4 per 100ml.

10. The number of samples per month required for detection of coliforms depends more on the geometric standard deviation than on the geometric mean density; more samples are required as the geometric standard deviation increases.

11. We recommend a minimum of six samples per month for microbiological monitoring of a water distribution system based on the following assumptions:

    a. A true arithmetic mean coliform density of $\leq$ 1 per 100ml indicates adequate reliability of protection.

    b. If a water system is found to violate the microbiological MCL more frequently than once a year, action will be taken to improve disinfection and/or other protection of the system.

    c. If a water system is found to violate the microbiological MCL once a year or less, no remedial action will be taken.

    d. The geometric standard deviation of the coliform distribution will be 100 or less.

12. Collection of two replicate 100 ml samples per site sampled increased the fraction of the sites where coliforms were found by about 50%. This suggests that increasing the volume of the standard sample to 200 ml would increase the detection of coliforms when they are present in densities of 5 to 10 per liter.

13. If the true arithmetic mean coliform density is $\geq$ 1 per 100ml and the geometric standard deviation is $\leq$ 100, violation of the microbiological MCL will occur twice a year if 6 samples per month are taken.

14. A water distribution system can be divided into sections which are hydraulically isolated; that is, coliforms may be present in one section and not be detected by samples collected from other sections. Significance differences in the frequency of coliform occurrence may sometimes be found among the isolated sections of one distribution system.

15. No significant differences were found in the frequency of coliform occurrence between peripheral and nonperipheral sampling locations. The frequency of coliform occurrence was not found to increase with distance from the water source. Thus, it is recommended that the sampling sites for a microbiological monitoring program be selected by a randomization procedure and that all parts of the distribution system be included.

SECTION 3

METHODS AND PROCEDURES

3.1 INTRODUCTION

In this section we define the variables accounted for within our sampling program and describe our protocols for sample collection, transit, analysis, quality assurance and data management. Information on the Woodbury Heights, New Jersey water distribution system and some of the data collected from that system are used to illustrate certain points. Information on and summaries of the data collected from the other systems are provided elsewhere including Sections 5 and 6 and the Appendices.

The Regulations require (U.S.E.P.A. 1976) that samples be collected at points representative of the conditions in the system. The term "representative" may be interpreted in several different ways, each of which has some validity depending upon the objectives of the monitoring program. One purpose of this project has been to define what "representative" should mean for water distribution systems in general, and to provide examples of how sampling programs can be designed to meet specific obectives. The sampling conducted for this project has had two major objectives: 1) to obtain random samples to provide data for fitting frequency distributions and 2) to test certain hypotheses about the sources, persistence, and transportation of bacteriological contamination of a water distribution system. These are not necessarily the objectives of a monitoring program, but the methodology used in the development of these programs can be applied to the development of monitoring programs once the objectives of such programs have been specifically defined. The objectives of a monitoring program to meet federal or state regulations are discussed in Section 1.

3.2 PLANNING A SAMPLING PROGRAM

We found it necessary to devise a method of designating exactly where a sample was collected and describing the real relationship among several samples in order to achieve the objectives of this project. This was accomplished by subdividing a distribution system into sections, locations, services and taps to form a hierarchy as described immediately hereafter. This method may have some utility for planning monitoring programs.

3.2.1 Hierarchy of Sampling Units

The physical units of sampling may be considered in a hierarchy from largest to smallest as: water system, distribution system, section, location, service, and tap. A water system as considered here meets the definition in

Section 141.2 (e) of the Regulations (U.S.E.P.A. 1976) and as such it includes "... any collection, treatment, storage, and distribution facilities ...." That part of the water system, after collection, treatment, and storage of water, that is used to distribute water is considered the distribution system.

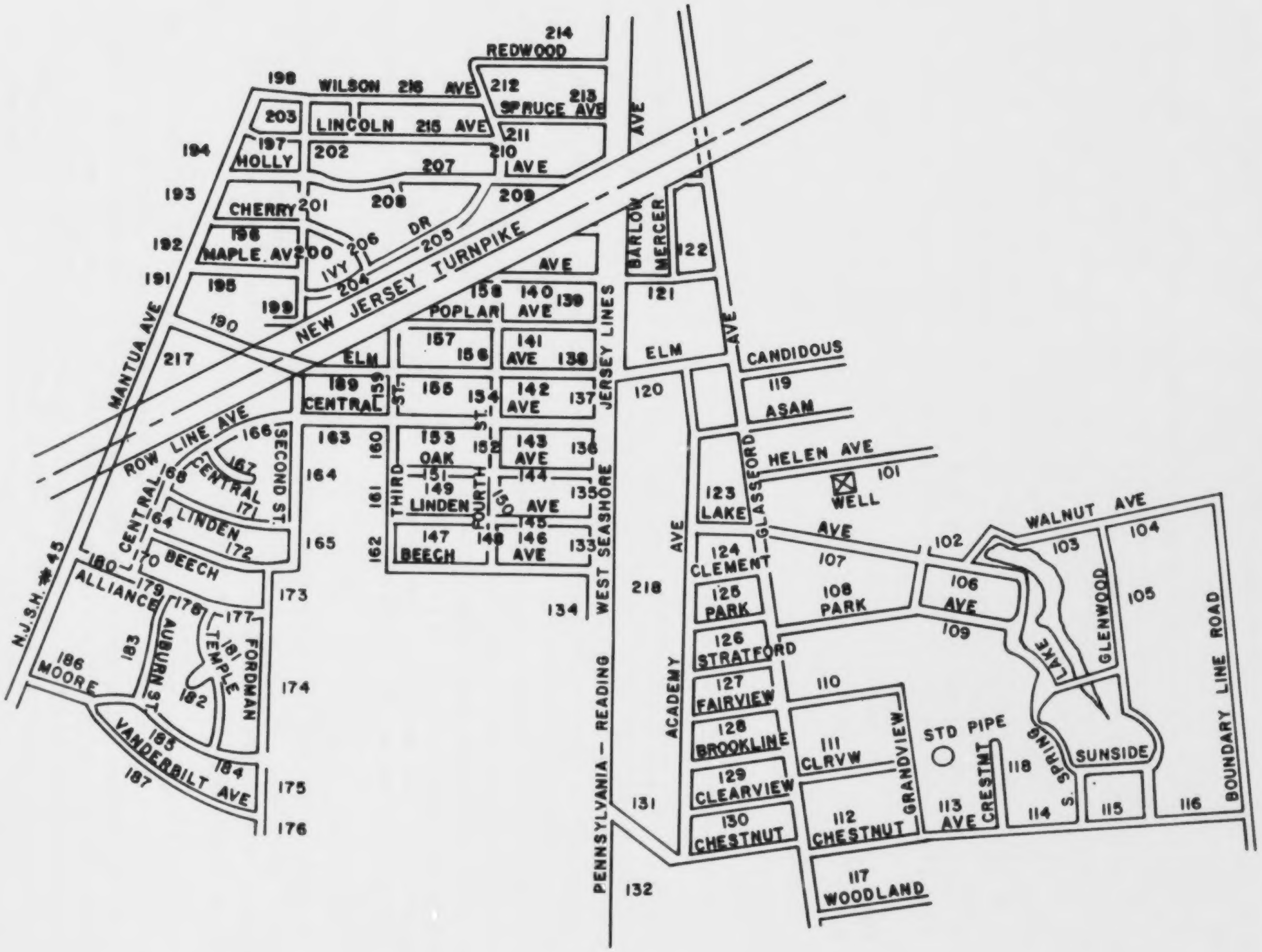
The definition of sections of a distribution system is more subjective. A section is defined as a portion in which the pattern of water flow is isolated to some degree from that of other sections or in which the inflow and outflow of water to and from the section are limited. The terms "to some degree" and "limited" are subjective. The desired result of delineating a section is to have a portion of the distribution system in which recognition of contamination within the section has potential importance to more than one location but less than the distribution system as a whole. A section is delineated from the patterns of mains and streets. We have assigned configuration types to sections, described in Section 6.

The pipes which make up a distribution system may be designated as transmission mains, distribution mains, and street laterals. Transmission mains convey water from the treatment facility to the sections of the distribution system and typically are the largest pipes of the system. A distribution main delivers water to street laterals within a section and to other sections. Street laterals deliver water to service connections and sometimes to other street laterals.

In this project, a sampling location was defined as a pipe between two connections supplying water to one or more service laterals. A location could be either a distribution main or a street lateral. Usually, a location corresponded to a city block with the main or lateral located under the street, between the connections under street intersections, and services on either side of the street receiving water from the same location. A joint in a distribution system is where three or more street laterals or distribution mains are joined and indicates where locations are separated. Occasionally, it was found that a street lateral served two or more blocks so that a sampling location was more than one block or that two street laterals served the opposite sides of the street in one block making that block two sampling locations.

A service is any building receiving water from the distribution system within a location. A tap is a faucet within a service from which water may be drawn. Fire hydrants represent another way of obtaining samples, but cannot be described simply as a service or tap.

The public water system of Woodbury Heights, New Jersey is used as an example of this hierarchial subdivision. Figure 3.1 is a map of the town. The numbers denote locations. Raw water is obtained and treated at a well at location 101 and stored in a standpipe between locations 111 and 112. Five sections have been delineated each containing between 16 and 31 locations. The sections are shown in Table 3.1. Each location contains at least 1 service, and a service may contain several taps. Thus, a location may be sampled at several discrete points within it.



NEW JERSEY TURNPIKE
MANTUA AVE
N.J.S.H. #45
WILSON 216 AVE
LINCOLN 215 AVE
REDWOOD
SPRUCE AVE
HOLLY
CHERRY
MAPLE AV
IVY
DR
AVE
POPLAR AVE
ELM
CENTRAL
ROW LINE AVE
SECOND ST.
THIRD ST.
FOURTH ST.
OAK
LINDEN
BEECH
ALLIANCE
AUBURN ST.
TEMPLE
FORDMAN
MOORE
VANDERBILT AVE
JERSEY LINES
WEST SEASHORE
PENNSYLVANIA – READING
BARLOW AVE
MERCER
ACADEMY AVE
GLASSFORD AVE
CANDIDOUS
ASAM
HELEN AVE
WELL
LAKE
CLEMENT
PARK
STRATFORD
FAIRVIEW
BROOKLINE
CLEARVIEW
CHESTNUT
CLRVW
GRANDVIEW
WOODLAND
STD PIPE
CRESTMT
S. SPRING
SUNSIDE
GLENWOOD
WALNUT AVE
BOUNDARY LINE ROAD
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218

Figure 3.1 Woodbury Heights Sampling Location Map

This hierarchy of sampling units is helpful for defining representative samples. Samples taken within the treatment facility cannot be considered representative of the water distribution system. Results from these samples only provide information on the bacteriological quality of the source water and the efficacy of treatment. This is but one possible source of contamination to the system as a whole.

TABLE 3.1

SECTIONS OF WOODBURY HEIGHTS, N.J.

Name	Locations included	No. of locations
East	101-119	22
East central	120-132, 218	16
West central	133-162	31
Southwest	163-188	26
North	189-217	30

Contamination may arise from any point in the distribution system and may occur in one section without occurring in the others. "Representative samples" should represent the different sections. If contamination is found in a section, follow up monitoring may focus on the section from which the contamination was found. The configurations of sections differ as described in Section 6. A particular configuration may be more prone to contamination than another. Information of this sort may be useful in not only designing a monitoring program, but also designing the distribution system itself.

The basic unit of sampling used for our study was the location. There are several advantages to this. The number of locations within a system is often large enough to be considered for application of random sampling, and yet small to allow for adequate data management. For very small systems, too few locations may be available for random sampling. Sampling of residential areas requires that someone be home and give permission for a sample to be collected. As a location often contains several services there is a better probability of obtaining a sample from a particular location than from a particular service. Also, multiple services may be sampled from a location if information on reproducibility is desired. Consideration must be given for check sampling, however. The service and tap sampled within a location must be noted if check sampling is to be conducted according to the current regulations (Section 141.21(d)(3 and 4)).

A service may have several taps from which samples may be obtained. The location of the tap may be important in consideration of subsequent contamination (Section 4). Also, only certain taps may be made accessible to the sampler. We have found it easier to gain permission to sample outside taps than those inside. From the point of view of monitoring the water in the distribution system per se, the tap location within a service should have no bearing aside from that of possible subsequent contamination. However, the tap location may be of importance if intraservice cross connections are being considered.

The designation of a section, location, or service is often based on the visible layout of the municipality rather than the pipe layout. Blueprints of pipe layout may be available, but their availability is sometimes limited. Maps of old distribution systems may be lost; new distribution and recently modified systems may lack adequate blueprints. One can use street layout as a first approximation of pipe layout. Exceptions exist, however, and there is no true substitute for accurate blueprints in designating the units of sampling.

Another consideration in sampling is time which can be considered in a hierarchial sense. The unit of time designated in the Regulations (U.S.E.P.A. 1976) is the month. In evaluating the sustained bacteriological quality of a system, however, time periods of many months may need to be considered. Seasonal changes in bacteriological quality may occur or the ability of a system to meet the coliform MCL's when contamination is continuous but at low densities or when sample size is small may best be evaluated over several months (Pipes and Christian 1978).

In theory sampling should occur randomly through time including night and day since water is supplied continuously and contimination may occur at any time. In practice this is not feasible. Systems requiring one or a few samples per month may be sampled only one or at most a few days. The times of the day or days of the week at which one has access to services or to personnel is limited. If a limited number of samples is taken and the samples are distributed over several days, this time may detract from efforts to obtain adequate sampling of different sections. The sampling pattern must be a compromise between evaluation of space and time differences in bacteriological water quality.

3.2.2 Random Sampling

At any time there is a very large (for all practical purposes, infinite) number of 100ml samples of water which could be collected from a water distribution system. An "ideal" experiment would be to drain the system instantaneously, examine each of the 100ml portions of water before the bacterial densities could change, and count each coliform organism in each sample. The results of this experiment would give the actual frequency distribution of coliform counts per 100ml. However, the water distribution system would be empty and the frequency distribution found would not necessarily describe the dispersion of coliform bacteria in the system once it were filled again.

The "ideal" experiment is clearly impossible; the frequency distribution describing the dispersion of coliforms in a water distribution system must be estimated from sampling data. The most important characteristic which the data used for fitting frequency distributions should have is that any 100ml portion of the water should have the same probability of being collected as any other 100ml portion of the water. This is the characteristic of a random sample but it is not possible, for practical reasons, to collect such a set of samples from a water distribution system. It is only possible to approximate random sampling and to understand the extent of deviation from randomness. The two variables considered in the randomization process are location and time.

At any given time, not all of the water in a distribution system is equally accessible. Samples can only be obtained from taps on the system. The sampling protocol calls for allowing the tap to run for at least 3 minutes to flush out the service lateral before the sample is taken. The sampling procedure is completed in less than 10 minutes. Thus, the part of the water which is accessible at any given time is the water which is within 3 to 10 minutes of time-of-flow from a tap.

Consider a residential service on block number 147 (Beech Avenue between Third Street and Fourth Street) on the sampling location map of Woodbury Heights, New Jersey (figure 3.1). The water main for that block is 6 inches in diameter, 700 feet long and contains approximately 1030 gallons of water or about 1.5 gallons per foot. There are seven houses on the north side of the street and eight on the south side or 15 potential sampling sites. If a service lateral is 50' of 3/4" pipe and in the house there are 30' of 1/2" pipe from the water meter to the tap to be sampled, then there is slightly more than one gallon of water between the street lateral and the tap. If the tap is allowed to run for 3 minutes at about 1 gallon per minute the service lateral and pipes in the building are flushed out three times but the water in the street lateral moves forward only about 2 feet (assuming no other water usage). If the tap is run for 10 minutes at one gallon per minute, the water moves forward in the main about 7 feet. Since there are, on the average 46.7 feet between service laterals, some fraction between 0.04 and 0.15 of the water in the main is accessible for sampling at a particular service. From this example it is clear that any tap on a particular service should be equivalent to any other tap on the same service in terms of the quality of water obtained for a sample but that each service on the block represents a different portion of the water in the main for sampling purposes. When this example is worked out for different pipe sizes, block lengths, service connections per block, etc. the numbers change somewhat, but in the context of what is considered to be acceptable, modern water supply practice the conclusions remain the same.

The best way to obtain a quasi-random sample would be to number each service connection on the system and select the service connections to be sampled from a table of random numbers. Unfortunately, this is impractical as mentioned previously because at the time of sampling there may be no one at home in a particular residence to provide access to a tap. Even when someone is at home, entrance to the residence for the purpose of sampling may be refused. Thus we adopted the street lateral as the location, numbering each as a potential sampling location and selecting location numbers for sampling from a table of random numbers. Then any service is considered to represent a sample from that street lateral. Even this approach does not always work out because it may not be possible to gain access to any of the services for sampling purposes.

In most cases (except for dead ends), it is expected that the water in a main will be flowing rather than still. Thus, a sample from a residence represents extracting a small fraction of the flowing water rather than extraction from a static body of water. Although there is very little longitudinal mixing of water flowing in a pipe, the fact that the water is flowing means that a larger fraction of the water in the system is potentially accessible for sampling.

Usually, we attempted to obtain samples from a tap in the yard because this avoided the need to enter a house. Of course, this was not possible in the winter and some home owners preferred that the samples be collected inside. The second most frequently sampled tap was the kitchen sink. Other taps which have been sampled include the bathroom lavatory, laundry room sink, and utility sink, (usually in the basement). Because the service lateral was flushed out several times before the samples were collected, there should have been no difference in water quality in samples collected at the different taps.

Since water is flowing in the distribution system, the coliform density in any part of the system may be expected to be changing continuously. To collect a random sample from the distribution system it would be necessary to randomize the time of sample collection as well as the locations of the samples. However, it would be futile to ring door bells at 3:00 am to ask permission to obtain water samples. As a practical matter, the time of sample collection for this project, as it would be for routine monitoring programs, was even more restricted by the necessity of processing the samples within a relatively short period of time after collection and the necessity of providing reasonable working hours for the samplers and laboratory personnel. Thus, most samples were collected on weekday mornings.

The bacteriological quality of water in a distribution system can vary with time. A major question for planning sampling programs is how rapidly and how often it changes. If the bacteriological quality of water in a distribution system remains stable for several days, data collected over that period can be lumped together for fitting frequency distributions. Changes in bacteriological quality would be recognized by significant changes in parameters such as fraction of sampling locations yielding coliforms, mean coliform density, etc. One of the results of this project has been information on temporal changes in the bacteriological quality of water in a distribution system and the reasonableness of calculating statistical parameters from data collected on several different days.

Data from random samples are needed for fitting to frequency distributions. It is impossible, in an absolute sense, to randomize collection of samples from a water distribution system because 1) not all of the water in the system is equally accessible at any time and 2) it is impractical to collect samples at certain times of the day. The best that can be done is to avoid known sources of bias and to test for changes in the parameters from day to day. This is more of a conceptual than a real problem. The objective of the project was to develop a model which can be used to specify the parameters of a routine monitoring program. As long as the data used for model development are comparable with data obtained from routine monitoring programs, the model should be valid.

3.2.3 Stratified Random Sampling

"Representative" sampling usually means selecting certain groups of samples to represent characteristics of the population being sampled. The division of the possible samples into groups is called "stratification".

The water distribution systems sampled have been stratified by dividing them into sections. Sections were selected insofar as possible to be isolated from other sections, but some of the systems sampled have had central portions which were not isolated, and the system LB had no isolated sections. The division of a large interconnected grid into sections has to be somewhat arbitrary and based on convenience of sampling.

The number of sections should be small in comparison with the number of samples collected during any time period. If the number of sections were equal to the number of samples collected, then there would be one sample from each section. This would be regular rather than random sampling. There have to be several samples per section for the requirements of a stratified random sampling pattern to be met.

The purpose of stratified random sampling in this project was to determine if there were differences in bacteriological quality among the different parts of a water distribution system. If such differences were found, a monitoring program would have to be designed somewhat differently from a program for a homogeneous system.

3.2.4 Other Sampling Patterns

One hypothesis which was proposed before the project started (Pipes and Christian, 1978) was that if there were a point source of bacteriological contamination such as occurs in the case of a cross connection, samples downstream of a positive sample would also be positive and all samples upstream of a negative sample would also be negative. Testing such a hypothesis required a linear sampling pattern along particular mains and the laboratory physical model.

Another hypothesis proposed before the project started (Pipes and Christian, 1978) is that the probability of finding coliform bacteria in a sample increases with distance from the water source and the greatest probability occurs at the periphery of the system. Testing the first part of the hypothesis required a sample stratified according to distance from the water source. Testing the second part of the hypothesis required a peripheral sampling pattern; that is, a set of samples collected at peripheral locations and a comparable set of samples collected at other locations. For the purpose of this project a peripheral location was defined as a street lateral which supplied water only to the service connections on that block. A dead end street lateral clearly meets this definition. A loop consists of two dead ends connected together; that is, two street laterals which, if not connected, would not influence the flow in any other parts of the system. Once these definitions are understood, it is clear that peripheral location can occur anywhere in the system, not just at the edges, and they can be close to the water source as well as distant from it.

3.2.5 Application to Monitoring Programs

It is unlikely that any program for routine monitoring of a water distribution system would be set up for random sampling. The randomization of sampling locations which was used for acquisition of some of our data sets approximated random sampling as closely as is possible given the physical limitations of a typical water distribution system. There is no reason to believe that these physical limitations seriously biased any data sets so they have been used for fitting frequency distributions. The frequency distributions selected have been used for evaluation of various types of sampling programs which might be used by water departments.

The term "representative locations" in the Regulations infers a stratified random sampling program. It is clear both from the Regulations and from our sampling experience that there should not be areas of a water distribution system which are never sampled. Recognition of sections will aid in preventing this. Most of the positive samples collected from system CV were from areas of the system which are never sampled by the Water Department samplers. The sampling locations should be changed from month to month so that all subsections of the system are covered over a period of time. Otherwise a serious problem in one part of the system could go undetected for long periods of time.

In order to cover all sections of a water distribution system it is necessary to obtain samples from private residences as well as from public and commercial buildings. Some water departments avoid sampling private residences. Our experience is that sampling private residences is slower than sampling public and commercial buildings because in many cases there is no one home and in some cases entrance to the residence is refused. However, it is possible to obtain large numbers of samples from private residences if adequate effort is made. Peripheral locations as well as other parts of the system should be sampled.

The goal of a bacteriological monitoring program should be to achieve the highest probability of detecting any bacterial contamination of the distribution system possible under the resource constraints of the program. This section has provided qualitative information on the design of a sampling program. Data and the results of tests of hypotheses about quantitative aspects are provided in Section 6.

3.3 COLLECTION AND ANALYSIS OF SAMPLES

For the development of procedures and protocols for this project, the primary references used were two EPA Manuals (Geldreich 1976 and Bordner, *et al.* 1978) and *Standard Methods*, 14th Ed. (APHA 1975). The methods of sampling and sample analysis were not major subjects of investigation. Well accepted procedures which have been developed and tested by others over a period of years were used. It would be a waste of space to summarize those procedures from the reference documents. Only a brief description of our specific adaptions of the accepted procedures which have been made is given here.

A Quality Assurance Manual has been written for the project giving detailed descriptions of the procedures being used for collection, transportation, handling, and analysis of samples. Copies of that Manual are available upon request.

3.3.1 Collection of Samples

Selection of location, service and tap have been described. Sampling involved recording information as to the address of the service, location of tap, time of sampling and sampler. The tap water was allowed to run for no less than three minutes at which time the temperature of the water and chlorine residual were taken. Chlorine measurements included free and total chlorine determined by the use of DPD kits (Hach Chemical Co.). Samples for bacteriological analyses were collected in sterile bottles containing thiosulfate for dechlorination. The number of bottles and volume of water per bottle varied according to desired replication and filtering volume. Volume ranged from 125mℓ to 3.8ℓ, and replicate bottle numbers ranged from 1 to 12 for normal sampling.

The temperature during transit to the laboratory should be < 10°C. Samples were stored in ice chests with frozen, sealed ice packs whenever necessary. The time from sampling to processing was less than 8 hr with the average time between 3 and 4 hr.

3.3.2 Analysis of Samples

The membrane filter procedure was used for coliform analysis of 100ml or 200ml on most samples. Each analyst filtered about 50ml of sterile buffered rinse water before and after each series of samples using the graduated cylinder, funnel assembly, and rinse bottle used for processing samples. The membrane from these rinse water controls were incubated on m-Endo medium in the normal manner and serve as negative controls. Serial dilutions of cultures of Escherichia coli (ATCC8739) and Enterobacter cloacae (ATCC13047) were also filtered as positive controls to demonstrate that the Endo medium produced the typical coliform colonies.

After portions from the samples and negative controls were filtered for coliform tests, one ml portions were removed for standard plate counts. A standard plate count was also made on the Endo medium and rinse water from all rinse bottles used that day to test for contamination. The Endo medium and the tryptone-glucose-extract agar was made up fresh for each sampling day. The pH and turbidity determinations were made on each sample after the portions had been removed for the coliform tests and standard plate counts.

In some instances a very turbid sample left a noticeable deposit of particulate matter on a membrane filter. Even though it is well known that such particulate matter can interfere with the development of coliform colonies, the filter was placed on Endo medium and incubated for 22-24 hours. If no coliform colonies developed, the filter was placed in lauryl tryptose broth and incubated for another 48 hrs. If gas was produced in the broth, this provided presumptive evidence of coliform bacteria. This, of course, does not give a coliform count for that sample, but the number of samples with coliforms present for any sampling trip is an important parameter to be evaluated accurately.

For parts of a distribution system where samples routinely contained large amounts of particulate matter, to preclude this interference from causing negative errors in the data generated, samples were filtered in several smaller aliquots to give a total of 100ml. If necessary, these filters were subjected to the non-quantitative procedure outlined above. Alternatively, larger volume samples were collected and a portion of the sample analyzed by the multiple tube fermentation test for comparison with the membrane filter results. The objective here was to assure that coliforms were detected if present and to obtain the best quantitative information possible, not to make a comparison of the efficiency of the two techniques.

Typical coliform colonies on Endo medium are dark with a green metallic sheen. Some colonies on Endo medium have dark centers with no sheen, and in some instances it was difficult to judge the presence or absence of the sheen. In those cases in which the presence or absence of the sheen was a difficult judgement, the colony was counted as a presumptive coliform and transferred to lauryl tryptose broth and brilliant green bile broth for verification. In this way, any doubt about the recognition of coliform colonies on Endo medium was resolved by testing rather than by opinion. A differential count of colonies with similar morphology was made, and coliform counts were adjusted according to the verification results. In addition all typical coliform colonies on a filter were picked for verification if there were five or less. If there were more than five coliform colonies on a filter, the number picked for verification depended on the total number of coliform colonies on all filters that required verification that day; however, the number was never less than five. Verification involved growth and gas production in lauryl tryptose broth (LTB) and brilliant green bile broth (BGLB), growth on EMB agar, growth and reactions on triple sugar iron agar (TSI), and Gram stain.

Dark colonies without a sheen were counted as atypical colonies. Also, red, and clear colonies were counted. A small number of these colonies were run through the verification procedure. One or more of the colonies from the filter with the ATCC organisms were also run through the verification procedure as a positive control.

Some fraction of the isolated verified coliforms and other isolated colonies from each system were identified using the API 20E system. Supplemental macrotube tests were employed when questions arose about identification. When few colonies were isolated in a system virtually all colonies were identified. When many colonies were collected the fraction was reduced to as low as one half to two thirds the isolates.

Water from each sample was used for standard plate count (SPC) determinations. Aliquots of 1 ml were used for pour plates with Trypticase Glucose Extract Agar (TGEA). Negative controls at the beginning and end of each flask of TGEA poured and triplicate samples every tenth sample were prepared for quality assurance. Plates were incubated at 35°C for 48 hour prior to counting.

The turbidity of each sample was determined with a Hach Model 2100A turbidimeter. Water was taken periodically for other chemical analyses. Chemical parameters included pH, alkalinity, hardness, and dissolved orthophosphate, ammonia, oxygen, nitrate, and nitrite concentrations. All procedures were in accordance with Standard Methods (APHA 1975) or the Manual of Methods for Chemical Analysis of Water and Wastes (U.S.E.P.A. 1974).

3.3.3 Special Quality Assurance Procedures

All laboratory personnel on this project kept individual logs in which all activities, such as preparation of a batch of medium or collection of samples and all data collected were recorded. Separate log books were maintained for the autoclave, hot air oven, pH meter, all incubators, media preparation, media controls, media performance, and laboratory water quality. Records of samples collected were kept on field sheets, and the data for each sampling site were recorded on permanent data sheets in a separate notebook. Each sample was identified by the sampling location number and replicate number. On the field sheets, the street address of the building sampled and the exact location on the tap used for sampling were recorded along with the water temperature, chlorine residual, and time of sampling. Thus, it was possible to resample the same tap if needed.

When the samples were being filtered, each analyst recorded, in chronological order, the identification numbers of the samples filtered. It was then possible to ascertain which samples were filtered before and after any given sample. When several positive samples were obtained on a sampling day, the records were checked to make sure that the positive samples occurred at random in the sampling order and among the analysts.

Records on all project activities were maintained at least in duplicate and sometimes in triplicate. It was possible to reconstruct the history of any sample, batch of medium, etc. from these records if the need arose. In this way, reliance upon the memories of project personnel was minimized and all data used for modeling was documented and could be subjected to critical scrutiny if necessary.

During the grant period we applied for certification equivalence through Region 3 of the U.S.E.P.A. After site visitation this equivalence was given.

3.3.4 Data Management

Beyond the maintenance of records described above, much of the information was stored in computer files. Two primary files used were the "location file", identifying the position of each location within the system and its pipeline characteristics, and the "bacteriological file", containing results from each sample taken. A representative "location file" and "bacteriological file" are shown in Figures 3.2 and 3.3, respectively. From these files most of our data analysis was done. We employed both the SAS programming facilities and Fortran IV for these analyses on the IBM 370 computer operated by UniColl.

```
                                   SAMPLE LOCATION FILE FOR WOODBURY HEIGHTS                        12:21 FRIDAY, AUGUST 7, 1981

OBS  BLOCKNUM  NUM_RES  CT_SERVS  PIPE_DIA  LENGTH  SOURCE1  SOURCE2  STREET      BETWEEN       AND_IN         PERIPH  SUB_AREA

  1    101        1         0       12.00    1200                     HELEN       GLASSBORO     GRANDVIEW                 E
  2    101N       0         0        8.00     800    101                          HELEN         CANDIDUS                  E
  3    101S       0         0       12.00    3100    101                          HELEN         WOODLAND                  E
  4    102        5         0        6.00     400    101S             LAKE        GRANDVIEW     WALNUT                    E
  5    103        5         0        6.00    1000    102              WALNUT      LAKE          GLENWOOD                  E
  6    104        5         0        6.00     800    103              WALNUT      BOUNDRY LINE  GLENWOOD         D        E
  7    105        5         0        6.00    1000    103              GLENWOOD    WALNUT        SUNSIDE                   E
  8    106        5         0        6.00     700    102      109     LAKE        WALNUT        PARK             L        E
  9    107        5         0        6.00    1100    101S             LAKE        GLASSBORO     GRANDVIEW                 E
 10    108        5         0        6.00     900    101S             PARK        GLASSBORO     GRANDVIEW        D        E
 11    109        5         0        6.00     700    101S     106     PARK        LAKE          GRANDVIEW        L        E
 12    110        5         0        6.00     800    101S             FAIRVIEW    GLASSBORO     GRANDVIEW        D        E
 13    111        5         0       12.00     800    101S             CLEARVIEW   GLASSBORO     GRANDVIEW        D        E
 14    112        5         0        6.00     800    101S             CHESTNUT    GLASSBORO     GRANDVIEW                 E
 15    113        5         0        6.00     600    101S             CHESTNUT    CRESTMONT     GRANDVIEW                 E
 16    114        5         0        6.00     600    113              CHESTNUT    CRESTMONT     S. SPRING                 E
 17    115        4         0        6.00     400    114              CHESTNUT    GLENWOOD      S. SPRING                 E
 18    116        4         0        6.00     600    115              CHESTNUT    GLENWOOD      BOUNDRY LINE     D        E
 19    117        5         0        6.00     800    101S             WOODLAND    EAST OF       GLASSBORO        D        E
 20    118        2         0        0.75     200    113              CRESTMONT   NORTH OF      CHESTNUT         D        E
 21    119E       5         0        8.00     550    101N             CANDIDUS    EAST OF       GLASSBORO                 E
 22    119W       5         0        6.00     550    101N             CANDIDUS    EAST OF       GLASSBORO                 E
 23    120N       0         5        4.00     400    119E             ELM         GLASSBORO     BARLOW           D        NC
 24    120S       0         5        4.00     400    119W             ELM         GLASSBORO     BARLOW                    NC
 25    121        5         1        6.00     600    119W             MAPLE       GLASSBORO     BARLOW           D        NC
 26    122        1         0        0.75     100    121              MERCER      NORTH OF      MAPLE            D        NC
 27    123        5         0        8.00    1100    101              LAKE        GLASSBORO     W. JERSEY                 EC
 28    123S       0         0        8.00     300    101      130     ACADEMY     LAKE          CHESTNUT                  EC
 29    124        5         0        6.00     300    123S             CLEMENT     GLASSBORO     ACADEMY          D        EC
 30    125        5         0        6.00     350    123S             PARK        GLASSBORO     ACADEMY          D        EC
```

Figure 3.2 Portion of Location File from Woodbury Heights, N.J.

```
------------------------------------------------ DATE=26APR79 ------------------------------------------------
      L S      R   S   A   V     T   P  P   A  A      R  C   C       P  P              L                    C  C
      O A      E   A   N   O     U   R  R   T  T      E  L   L       L  L    C   C     O                    O  O
      C M      P   M   A   L   T R   E  E   Y  Y  R   D  E   R       A  T    L   L     C                    N  N
      _ P      _   P   L   _   E B   _  _   P  _  E   _  A   _       T  _    2   2     _                    _  _
      T L      S   _   _   F   M _   C  C   I  C  D   C  R   C       E  C    _   _  S  S                    C  C
 O  S Y E      A   V   R   I   P N   O  O   C  O  _   O  _   O       _  O    T   F  I  A                    O  O
 B  U P _      M   C   E   L   _ T   L  D   A  D  C   D  C   D       C  D    O   R  T  M                    L  D
 S  B E C      P   L   P   T   C U   I  E   L  E  O   E  O   E       O  E    T   E  E  P                    I  E
13  E R 113    2  125  1  100  . .   0      0     0      0         0.0     0.30  .  O  811 CHESTNUT         0
14  E R 113    3  125  1  100  . .   0      0     0      0         0.0     0.30  .  O  811 CHESTNUT         0
15  B C 120M   2  125  1  100  . .   0      0     1      0         0.0     0.15  .  B  MUNICIPAL BLDG       0
16  B C 120M   3  125  1  100  . .   0      0     0      0         1.5     0.15  .  B  MUNICIPAL BLDG       0
17  B R 121C   2  125  1  100  . .   0      0     0      0         1.0     0.00  .  O  534 MAPLE            0
18  B R 121C   3  125  1  100  . .   1      0     0      0         0.0     0.00  .  O  534 MAPLE            1
19  B C 121W   2  125  1  100  . 1   0      0     0      0         0.0     0.00  0  I  CAR WASH MPLE        0
20  B C 121W   3  125  1  100  . 1   0      0     0      0         0.0     0.00  0  I  CAR WASH MPLE        0
21  S R 170    2  125  1  100  . .  11      6     .  TN  0        79.5     0.60  .  O  105 CENTRAL         11
22  S R 170    3  125  1  100  . .  16      . TN  .  TN  0        35.0     0.60  .  O  105 CENTRAL         16
23  S C 180AK  2  125  1  100  . .   0      0     0      0         0.0     0.60  .  K  PERKIN'S REST.       0
24  S C 180AK  3  125  1  100  . .   0      0     0      0         0.0     0.60  .  K  PERKIN'S REST.       0
25  S C 180AM  2  125  1  100  . .   0      0     0      1         0.0     0.50  .  B  PERKIN'S REST.       0
26  S C 180AM  3  125  1  100  . .  17     24     .  TN  .  TN   151.5     0.50  .  B  PERKIN'S REST.      13
27  S R 183    2  125  1  100  . .   0      0     0      1         0.0     0.10  .  O  424 AUBURN           0
28  S R 183    3  125  1  100  . .   0      0     0      0       101.0     0.10  .  O  424 AUBURN           0
29  S C 188M   2  125  1  100  . .   0      0     0      0        15.0     0.40  .  B  WOOLCO MOORE+45      0
30  S C 188M   3  125  1  100  . .   0      0     0      0         0.0     0.40  .  B  WOOLCO MOORE+45      0
31  S C 188S   2  125  1  100  . .   0      0     0      0         0.0     0.00  .  R  WOOLCO MOORE+45      0
32  S C 188S   3  125  1  100  . .   0      0     0      0         0.0     0.00  .  R  WOOLCO MOORE+45      0
```

Figure 3.3 Portion of Bacteriological File from Woodbury Heights, N.J.

SECTION 4

SAMPLE VALIDITY

4.1 INTRODUCTION

For the purposes of this project, a valid result from a sample is one in which the coliform count on the membrane filter accurately reflects the coliform density in the main or street lateral supplying the service sampled. If it is assumed that the coliforms suspended in the water in the main or street lateral are carried through the service pipes and out the tap, there are four different occurrences which can invalidate the results.

Coliforms may be introduced into a sample by non-sterile sample bottles, washings from the surface of the tap sampled, transfer from the hands of the sampler, leakage of contaminated water into the sample bottle during transport, or from the glassware used for measuring 100ml of the sample and containing it during filtration. We use the term "subsequent contamination" to describe all these cases in which coliforms could be introduced into the water subsequent to its discharge from the service pipes. Our studies of these occurrences and the steps we took to prevent them are described in Section 4.2.

Coliform bacteria are living organisms. As such they may die or they may reproduce. Either death or reproduction of coliforms in a water sample can occur between the time the sample is collected and when the sample is filtered. Changes in coliform density due either to death or reproduction would produce inaccurate results. Our studies of these changes are described in Section 4.3.

Certain types of particulate matter in the samples or large numbers of non-coliform bacteria may prevent the growth of coliform bacteria on the surface of membrane filters to form typical coliform colonies on M-Endo medium. This project did not include an investigation of these types of interferences as such. However, evidence of interference due to particulate matter in a few samples was observed. This is described in Section 4.4 along with the procedures used to identify its occurrence.

Non-coliform bacteria in the water sample may grow on M-Endo medium and produce colonies with a green sheen or coliform bacteria may produce colonies without a sheen. These two occurrences are known respectively as "false positive results" and "false negative results." False positive results are eliminated by the confirmation procedure. False negative results are more difficult to deal with in a routine monitoring program. Our experiences with both false positives and false negatives are described in Section 4.5.

4.2 SUBSEQUENT CONTAMINATION

When faced with a set of coliform data consisting of many zeros and a few counts such as 31, 2, 17, 1 etc., the person responsible for the system is often inclined to believe the zeros and to blame the positive results, at least the high positive results, on bad samples. Excuses such as "the lab sent over some bad sample bottles," "the tap was dirty," and "the sampler stuck his finger in the bottle" are frequently heard, particularly if check samples were negative. This type of thinking can lead to indications of failure of the disinfection process being ignored; however, it must be admitted that such mistakes can occur and steps should be taken to avoid them.

4.2.1. Sterility of Sample Bottles

Several steps were taken to assure the sterility of sample bottles used in this project. A log of the temperature and time of autoclaving for each use of the autoclave was maintained. Some sample bottles in each batch autoclaved were labeled with indicator tape on which the word "STERILE" appeared when there was adequate exposure to sterilizing temperature. Every three months the effectiveness of the autoclave was checked by autoclaving a strip of *Bacillus* spores and then incubating it in trypticase soy broth to obtain evidence that all of the spores were killed. With each batch of sample bottles autoclaved, 10ml of trypticase soy broth made up with distilled water containing at least 300 viable bacteria per ml was autoclaved in a culture tube in a 500ml sample bottle. The sample bottle was then incubated for three days and observed for growth.

In seven of the community systems sampled, the finished water from the well or treatment plant had many fewer positive coliform samples and lower coliform densities than the water collected from the distribution system. Well water samples were not obtained for systems MW and BG. These data are presented in Table 4.1 and are evidence that the sample bottles used in this project were properly sterilized. The data in Table 4.1 include all samples and the volumes filtered varied depending upon the system and sampling location.

4.2.2 Tap Location

The Regulations state that samples for the determination of microbiological water quality "shall be taken at points which are representative of the conditions within the distribution system." (U.S.E.P.A. 1976). This may be interpreted in different ways (Geldreich 1971). Samples may be taken from different sampling sites including private residences, public buildings such as schools and fire houses, commercial establishments, and fire hydrants. Also, samples may be taken from bathrooms, kitchens, utility rooms, outside hydrants, or drinking fountains. Interviews of personnel of water departments indicated that some communities sample solely from residences, others from fire hydrants, others from public buildings and commercial establishments, and others from combinations of location types. Public restrooms by their very nature are used by a diversity of people with an equally diverse assortment of personal cleanliness habits, and taps in public restrooms may be subjected to contamination not common to other faucets. Also,

TABLE 4.1

COMPARISON OF POSITIVE COLIFORM RESULTS IN SOURCE WATER AND DISTRIBUTION SYSTEM SAMPLES FOR SEVEN SYSTEMS

System	Sampling Period	Raw Water Samples			Finished Water Samples			Distribution System Samples		
		Number	Fraction Positive	Density[a] (per 100ml)	Number	Fraction Positive	Density (per 100ml)	Number	Fraction Positive	Density[a] (per 100ml)
CV	Jan.-Feb.1979	10	1.0	TNTC	39	0	0	252	0.020	0.058
	March 1979				24	0	0	414	0.027	>0.069
WH	April-May 1979	18	0.111	0.025	5	0	0	227	0.075	>0.177
	May 1979	101	0.010	0.010	96	0.031	0.073	298	0.258	1.688
	May-June 1981	29	0	0	58	0	0	342	0.108	>1.278
LB-BB	June 1979	34	0.029	0.029	21	0	0	213	0.061	0.441
	July 1979	11	0	0	16	0	0	198	0.136	>9.25
LB-TP	June 1979	28	0	0	18	0	0	192	0.276	>3.29
	July 1979	30	0	0	10	0	0	309	0.233	>5.53
DT	Dec.-Jan. 1980	15	1.000	> 693	49	0	0	242	0.025	0.028
	Jan.-Feb. 1980	12	1.000	900	55	0	0	222	0.032	>0.429
BL	March 1980	35	0.029	0.029	135	0	0	260	0.019	0.019
	April 1980	0	-	-	0	-	-	319	0.034	>0.539
	June 1981	25	0	0	30	0.033	0.033	299	0.094	0.712
MI[b]	May-June 1980	52	0					412	0.121	>0.396
SR[c]	Oct.-Dec. 1980				84	0.012	0.012	270	0.107	1.19

Notes:
- a. The > symbol indicates that at least one sample had a count of >80 per 100ml (TNTC).
- b. System MI used unchlorinated well water, so no finished water samples could be collected.
- c. System SR used chlorinated well water, but there was no point of access to the well water before chlorination.

TABLE 4.2.

COMPARISON OF COLIFORM RESULTS FROM DIFFERENT SAMPLING SITES

System	Sampling Period	Private Residences			Public Restrooms			Commercial Establishments Municipal Buildings		
		No. of Samples	Fraction Positive	Mean Density (per 100ml)	No. of Samples	Fraction Positive	Mean Density (per 100ml)	No. of Samples	Fraction Positive	Mean Density (per 100ml)
CF	Feb.1979	172	0.029	0.302	32	0	0	28	0	0
	March 1979	268	0.019	0.052	6	0.17	0.17	8	0	0
WH	April-May 1979	114	0.070	>1.070	20	0.05	0.65	19	0	0
	May 1979	274	0.263	1.657	10	0.10	0.10	17	0.176	0.647
	May-June 1979									
LB-BB	June 1979	197	0.081	0.523	2	0	0	2	0	0
	July 1979	174	0.218	>9.184	0	-	-	24	0.333	2.792
LB-TP	June 1979	226	0.292	>2.774	4	0	0	12	0.083	0.250
	July 1979	252	0.230	>6.798	8	0	0	49	0.082	>1.837
DT	Dec.-Jan.1980	207	0.024	0.034	17	0.059	0.059	18	0	0
	Jan.-Feb.1980	136	0.036	>0.728	16	0.062	0.125	16	0	0
BL	March 1980	215	0.009	0.009	25	0.040	0.040	20	0.050	0.050
	April 1980	261	0.031	>0.839	8	0	0	50	0.120	0.180
	June 1981									

contamination of faucets can result from aerosols produced by toilet flushing (Gerba et al. 1975). Thus it was postulated that samples from public restrooms may provide a larger fraction of positive coliform results than those from other taps due to subsequent contamination.

Two approaches were taken to examine the influence of faucet type on subsequent contamination. Samples were obtained from public restrooms and from other taps in commercial establishments in six of the systems. These data were analyzed with respect to the probability of positive results from each of the faucet types. In Table 4.2 the results of sampling of private residences, public restroom faucets, and taps in other commercial or municipal buildings for six water distribution systems are shown. Systems MI, SR, MW, and BG were small residential communities with no municipal or commercial buildings available for sampling. The data show no differences between samples collected in public restrooms and those collected in private residences or those collected from other taps in municipal and commercial buildings.

The second approach was used to eliminate any space dependent variations in coliform density. For this approach two faucets from either a commercial establishment or public building were sampled within 15 minutes of one another. A restroom faucet was sampled as well as another faucet within the same building. The results of this study are shown in Table 4.3. One positive sample was obtained from the 17 sites sampled. The positive sample was from a restroom faucet. These data suggest the possibility of subsequent contamination from restroom faucets, although they do not prove that it occurred.

TABLE 4.3

COMPARISON OF PAIRED SAMPLES FROM COMMERCIAL ESTABLISHMENTS

System	Number of sites sampled	Vol. filtered	Public Restroom		Other Faucet
			Number Positive	Coliform Count	Number Positive
CV	8	100ml	0		0
		800ml	0		0
		100ml	0		0
WH	9	100ml	0		0
		100ml	1	(13)	0
Totals	17		1		0
Fraction Positive			0.059		0.0
95% Confidence Interval			0.003 to 0.254		0.0 to 0.167

4.2.3 Sampler Error

Water flowing into the bottle may accidentally come in contact with the sampler's finger, or the sampler's finger may brush against the lip of the sample bottle. The appropriate course of action should be to eliminate that sample immediately. However, often samplers have a set number of bottles for an equal number of samples to be taken on the day or may not be aware of the incident. Also, during the processing of samples, contamination of the sample by the finger of the analyst may occur. If no coliforms are present upon testing, these errors are likely to be forgotten. If, however, a positive coliform result is obtained, these errors may be used as an explanation by the sampler. This situation poses a problem in interpreting the validity of positive coliform results.

Three questions were asked in this study. 1. If a finger were placed into a water sample, what is the probability of obtaining a positive coliform result? 2. If a positive coliform result is obtained, is the number of coliforms always large enough as to suggest this route of subsequent contamination? 3. When positive coliform results are obtained, are the species of organisms different from those normally found in distribution systems?

One hundred ml of dechlorinated tap water were placed in 125ml sample bottles, sterilized by autoclaving at 121°C for 30 min., and cooled to room temperature. Test subjects were asked to place one finger into the mouth of the bottle and shake the bottle vigorously for approximately 5 sec. The bottle was then capped, returned to the laboratory, and processed by membrane filtration techniques. Selected colonies from positive plates were isolated, verified, and identified by the normal procedures used for this project.

Test subjects were divided into three categories: 1. Water samplers from our laboratory, 2. Water samplers and water department employees from four different municipalities, and 3. Drexel University staff and students. In categories 1. and 3., a portion of the samples were taken from the same subjects on different days. In category 2. each sample represented a different subject.

The procedure used to determine subsequent contamination from fingers was designed to maximize the probability of contamination. Thus the results presented may be interpreted as a worst case situation. Despite this approach only 5.4% of the samples contained coliforms (Table 4.4). Positive coliform results were not found from any municipal water department personnel. The positive results from Drexel University water samplers were more numerous than for the other groups. These water samplers were also responsible for transfering coliform cultures and bacterial identifications. It appears that a positive coliform result may occur from the hands of the sampler, but such results are not highly probable.

Initially it was postulated that if a positive result were obtained, the potential for this route of contamination may be recognized. A water sample contaminated from a finger may have an unusually high bacterial density or a species of bacteria not normally found in distribution systems. However

TABLE 4.4

TESTS FOR SUBSEQUENT CONTAMINATION FROM FINGERS OF SAMPLERS

Subjects	No. of Samples Taken	No. with Coliforms	Fraction Contaminated (95% confidence interval)
Project staff	12	1	0.083 (0.000-0.236)
Water Department Employees	21	0	0 (0.000-0.087)
Drexel Staff and Students	78	5	0.064 (0.010-0.119)
Overall	111	6	0.054 (0.010-0.096)

when positive samples were obtained the number of coliforms or non-coliforms ranged from 1 to too numerous to count (TNTC = >80/100ml). Thus no indication of subsequent contamination could be found based merely on the number of colonies present. Also, the species of organisms found were similar to those we have found in water distribution systems. Coliforms found included *Klebsiella pneumoniae*, *Enterobacter cloacae*, and *Enterobacter agglomerans* as determined using the API system. The possibility of this type of subsequent contamination could not be determined by species identification.

4.2.4 Contamination During Transit In Ice

It is recommended that microbiological samples be stored at less than 4°C during transit to the laboratory (Bordner et al 1978). This may be done by keeping samples in a chest filled with ice. As the ice melts the sample bottles may have their tops immersed in the ice water. This immersion may be a route of subsequent contamination. As the bottles cool, differences in the coefficients of expansion of the materials used for the bottles and the caps may cause the caps to loosen. As cooling occurs the air within the bottle may contract producing a negative pressure. A combination of these two events may allow the intrusion of the surrounding ice water into the bottle and thus subsequent contamination. Previously, Geldreich (1975) has cautioned about this route of contamination but did not provide quantitation. A series of experiments to determine the likelihood of occurrence of this route of subsequent contamination was conducted.

Sample bottles were filled with phosphate buffer made up with distilled deionized water and autoclaved at 121°C for 30 min. Caps were tightly secured after cooling of the water. The bottles were then placed into ice water for 4 hours. The time in which the bottle tops were immersed in the ice water was controlled. In some cases the tops of the bottles were immersed for the entire 4 hours. Also, tops were immersed for 0.5 hr and 0.08 hr after cooling for 3.5 hr and 3.92 hr, respectively. One hundred ml of water from each bottle was then membrane filtered, and the filter was incubated on Endo medium for 22 hr at 35°C. Both coliform and non-coliform colonies were counted. Also, the ice water was sampled for coliforms by the same membrane filter technique.

Four different types of bottles were employed: 1. 125 ml glass dilution bottles with plastic caps, 2. 1 qt. (945 ml) glass bottles with plastic caps, 3. 125 ml polypropylene bottles with polypropylene caps, 4. 1 liter polypropylene bottles with polypropylene caps. In one experiment surface water was added to the ice water to increase the coliform density.

Bottles were immersed in ice water containing different densities of coliforms ranging from 1 CFU/100ml to 10^5 CFU/100ml. The results in Table 4.5 represent a composite of all bottle types and all times of immersion. Subsequent contamination occurred at all coliform densities tested. At low densities (< 10 CFU/100ml) as many as 12.5% of the bottles were positive for coliforms. At a high coliform density (10^5 CFU/100ml due to the introduction of creek water) all bottles were positive for coliforms. Unless one is assured that ice used in transit is completely devoid of coliforms, the potential for contamination is too great to allow the transit of sample bottles in ice where there is a possibility of bottle tops being immersed in the ice water.

TABLE 4.5

CONTAMINATION OF SAMPLE BOTTLES IMMERSED IN ICE WATER CONTAINING DIFFERENT DENSITIES OF COLIFORMS

Density of Coliforms in Ice Water (per 100 ml)	No. of Bottles Under Ice Water	No. of Bottles With Coliforms	Fraction Positive	(95% Confidence Interval)
< 1	17	1	0.050	(0.003-0.254)
~ 2	24	3	0.125	(0.035-0.308)
~ 10	10	1	0.100	(0.005-0.397)
~ 10^5	48	48	1.000	(0.925-1.000)

Ice may contain bacteria which are not coliforms but which may grow on Endo medium. Such organisms may also contaminate samples and may be responsible for interference with coliform counting (Geldreich et al 1978). We were able to use this population of organisms as a tracer in our subsequent contamination study. The ice we used could be divided into two states with respect to non-coliform organism densities. "Uncontaminated" ice water possessed < 1 CFU/100ml of these organisms, and "contaminated" ice water contained > 200 CFU/100ml. Experiments were divided according to these criteria. In Table 4.6 the results of experiments with different periods of immersion of bottle tops are shown. In "contaminated" ice water positive results were obtained in all but one case (65 of 66 bottles were positive). Both glass and polypropylene bottles could be contaminated, and contamination occurred even when bottles were immersed for as little as 5 min. In "uncontaminated" ice water the fraction positive was considerably less (2 of 17 bottles were positive). The positive results in this case were for the larger sample bottles. The results confirm those for coliforms and extend the observation that only a short period of immersion is required to obtain contamination.

Based on these and other findings, it is clear that bottles should not be stored in ice in such a way as to allow immersion of the bottle top (Geldreich 1975). However, samples should be maintained at 1-4°C during transit whenever possible. This problem may be alleviated by using sealed coolant instead of loose ice.

4.2.5 Carry Over During Filtration

In the membrane filter technique for coliforms the same filtration unit is used for a number of samples. It is recommended that a filtration unit be rinsed between each sample to insure that all coliform bacteria in the sample reach the filter and to avoid the carry over of coliforms from one sample to the next. If coliforms were found in high density in one sample, inadequate rinsing could permit coliforms to be found in one or more subsequent samples. This source of subsequent contamination must be avoided or at least identified if it were to happen. The prescribed technique of rinsing is to "rinse the sides of the funnel at least twice with 200 ml of sterile dilution water". However, a number of variables in procedure exist which are not addressed in this general protocol. These variables include 1) measuring sample volume in graduated cylinders or directly in the funnel, 2) rinsing with a squeeze bottle of dilution water or from a graduated cylinder, and 3) using different brands of filter funnels. These variables were examined to develop a rinsing procedure to minimize the potential for carry over.

The potential for carry over was investigated by pure culture experiments and observations of filtering order of field samples. Pure cultures of *Escherichia coli* or *Enterobacter cloacae* were grown in nutrient broth overnight and diluted in sterile buffered dilution water to appropriate densities. One hundred ml of diluted cultures were measured in graduate cylinders and then filtered in either glass or plastic polysulfone filter funnels. The glass funnel was held in place by a spring loaded clamp while the plastic funnel was held in place by a magnetic ring. Two rinsing protocols were followed. For one series of filters, a 30 ml aliquot of sterile dilution water was placed in the graduated cylinder used previously, and this aliquot was used to wash down the sides of the filter funnel while vacuum was still being applied. This procedure was conducted a total of three times. The filter was then removed and incubated on Endo medium. A second series of rinsing was done on replicate samples using sterile dilution

TABLE 4.6

POSITIVE SAMPLES FROM DIFFERENT TYPES OF BOTTLES IMMERSED IN ICE WATER FOR VARIOUS PERIODS OF TIME

Bottle Type	Hrs in Ice Water	No. of Bottles	No. of Bottles With Positive Results	Fraction	(95% Confidence Interval)
Contaminated ice water (200 CFU/110 ml)					
125 ml glass	4	22	21	0.95	(0.863-1.000)
	0.5	12	12	1.00	(0.816-1.000)
	0.08	8	8	1.00	(0.685-1.000)
125 ml poly.	4	8	8	1.00	(0.685-1.000)
	0.5	8	8	1.00	(0.685-1.000)
	0.08	8	8	1.00	(0.685-1.000)
Uncontaminated ice water (1 CFU/100 ml)					
125 ml glass	4	4	0	0	(0.000-0.527)
125 ml poly.	4	5	0	0	(0.000-0.500)
945 ml glass	4	4	1	0.25	(0.013-0.751)
1 poly.	4	4	1	0.25	(0.013-0.751)

water kept in a squeeze bottle. In this series the graduated cylinder was rinsed well with a jet of water from the squeeze bottle and poured onto the filter. Approximately 5 to 10 ml of water were used for this process. The sides of the filter funnel were washed with a jet of water from the squeeze bottle (20-30 ml). Again the filter was removed and incubated on Endo medium. The bottom ring of the plastic filter funnel top was then rinsed with a jet of sterile dilution water (5-10 ml). For both rinse series a new filter was placed in the filter funnel and 100ml of sterile dilution water was filtered. These filters were incubated on Endo medium to determine the extent of carry over.

The results of one series of laboratory experiments using E. cloacae are shown in Table 4.7. The variables investigated in this study included coliform density, filter funnel type, and rinsing protocol. In this study neither rinsing protocol nor filter funnel type altered the potential for carry over. However, the initial density of coliforms was a major determinant for the potential of carry over. Little or no carry over was found at the lower two densities studied, and carry over was almost unavoidable at the higher two densities. In other experiments with E. coli and contaminated surface water, carry over routinely occurred after densities too numerous to count (>80 coliforms/100 ml) were filtered. During these experiments squeeze bottle rinsing was found to be the more effective protocol. Also, the importance of rinsing the bottom ring of the plastic filter funnel should be noted. Apparently the magnetic ring which holds the filter funnel in place does not produce enough pressure to prevent seepage of water out of the funnel during filteration. Extra rinsing of this ring to help eliminate carry over is recommended.

Samples collected in the field were used to demonstrate the reliability of the squeeze bottle procedure. Duplicate 100 ml samples were taken at all locations sampled, and occasionally 1 liter samples were taken in addition. Filtration of samples were carried out 2 to 6 hours later. Samples were filtered, filter funnels were rinsed by the squeeze bottle protocol, and the order of filtering recorded. The extent of possible carry over was determined by examination of sequential positive coliform results. Carry over was considered possible if two criteria were fulfilled: 1) if the positive coliform result immediately followed in order of filtration another positive result with a greater coliform density, and 2) if the positive coliform result was not supported by a positive result in a replicate sample.

The coliform results from selected field samples are presented in Table 4.8. One hundred three samples gave positive results during this study of which 4 were considered possible cases of carry over by the two criteria described above. Two of these were from 1 liter samples in which a count of >80/100 ml was followed by a count of 14/100 ml and then 7/100ml. The next filtered sample was not positive for coliforms. One liter samples are rarely used in routine analyses, but may offer a greater possibility of carry over than 100 ml samples. In liter samples the 300 ml filter funnel is filled to the top and water is in contact with the funnel for a longer period of time. The remaining two possible carry over cases each gave 1/100 ml after a sample containing >10/100 ml was filtered. As stated, the replicates of these samples were negative. However, at low coliform densities it

TABLE 4.7

RELATIONSHIPS OF COLIFORM DENSITY, TYPE OF FILTER FUNNEL AND RINSING PROTOCOL AS DETERMINANTS OF CARRY OVER

	Coliform Count on Each Second Filter Coliform Density ± (SD) (per 100ml)				
	Glass Funnel		Plastic Filter Funnel		
Coliform Density of First Sample Filtered (per 100ml)	Graduated Cylinder Rinse	Squeeze Bottle Rinse	Graduated Cylinder Rinse	Squeeze Bottle Rinse	Fraction of Second Filters Positive (95% Confidence Interval) Coliform Density (95% Confidence Interval)
30,000	3,11,9 7.7 ± (4.2)	8,5,6 6.3 ± (1.5)	7,25,16 16.0 ± (9.0)	9,7,3 6.3 ± (3.1)	1.000 (0.764-1.000) 9.1 (5.2-13.0)
3,000	0,0,1 0.3 ± (0.6)	3,1 2.0 ± (1.41)	1,3,2 2.0 ± (1.0)	0,1,0 6.3 ± (3.1)	0.635 (0.333-0.865) 1.1 (0.4-1.8)
125	0,0,0, 0.0 ± (0.0)	0,0,0 0.0 ± (0.0)	0,0,1 0.3 ± (0.6)	0,1,0 0.0 ± (0.0)	0.083 (0.004-0.346) 0.1 (0.0-0.3)
19	0,0 0.0 ± (0.0)	0,0,0 0.0 ± (0.0)	0,0,0 0.0 ± (0.0)	0,0,0 0.0 ± (0.0)	0.000 (0.000-0.250) 0.0

is to be expected that some replicate samples from a site will be positive when the other is negative.

TABLE 4.8

POTENTIAL CARRY OVER ANALYSIS OF SELECTED SAMPLES COLLECTED IN THE FIELD

Initial Coliforms Count	Coliform Count of Next Sample (No. of Second Samples with the Density Listed)
1	0(30), 1(6), 3(3), 4(2), 5(1), 11(1), 15(1), 35(1), 48(1), TNTC(1)
2	0(9), 2(2), 11(1)
3	0(8), 1(3), 3(1), 9(1)
4	0(3)
5	0(1), 24(1)
6-10	0(7), 1(1)
11-20	0(9), 7(1*)
21-40	0(2), 1(1*,1), 2(1)
41-60	0(2), 1(1)
61-80	1(1*)
> 80	14(1*)

* potential carry over

Based on these findings about subsequent contamination, carry over is a source of error. Therefore, to avoid carry over or recognize it when it does occur, the following procedures are recommended:

1. Rinsing of any graduated cylinder used to determine sample volume must be included in the protocol.
2. Rinsing with a jet of water from a squeeze bottle is at least as efficient and perhaps more efficient than rinsing from a graduated cylinder.
3. The bottom ring of the plastic filter funnel top should be rinsed between each sample.
4. The order of filtration must be recorded and positive results interpreted within the context of filtration order.
5. Any positive results filtered following a positive result of greater than 20 coliforms/100ml should be considered suspect of subsequent contamination by carry over.

4.2.6 Summary on Subsequent Contamination

To assess accurately the microbiological water quality of a water distribution system, procedures must be designed to avoid subsequent contamination or account for it if it does occur. Potential routes of subsequent contamination include 1) tap location, 2) sampler error by placing his finger in the bottle, 3) transit of sample bottles in ice, and 4) carry over during filtration. Of these, transit in ice and carry over during filtration give the greatest likelihood of subsequent contamination. Sampling public restroom faucets and finger contamination of samples demonstrated a low probability of leading to subsequent contaminated positive results. However, any chance of subsequent contamination should be minimized or eliminated if possible.

Based on observations from these studies, recommendations for sampling and sample processing are as follows:

1. Routine sampling should avoid public restroom faucets. If a sample from a commercial establishment or public building is desired, the sample should be taken from a faucet from a kitchen, utility room or other non-restroom location.

2. If a water sample comes in contact with a finger or part of the sampler's hand at any time, the sample should be discarded. Although a negative result is highly probable, the chance for a positive result exists. However, the probability of obtaining a positive result is so low that the use of this as an excuse for a positive result is tenuous at best.

3. Water samples should be cooled to 4°C during transit but the tops of sample bottles should never be allowed to be immersed in ice or ice water. Sealed containers of ice or coolant should always be used in preference to loose ice. This precaution is so simple that no substitute short of portable refrigeration should be allowed.

4. Carry over of coliforms from one sample to the next during filtration is of major concern as a route of subsequent contamination. If a sample contains a large number of coliforms (>80 coliforms/100 ml), subsequent contamination is highly probable. Optimum rinsing procedures must be maintained at all times. Rinsing by squeeze bottle is at least as effective and probably more effective than using a graduated cylinder rinse. Any graduated cylinder used to measure sample volume must be rinsed into the sample filter. The bottom ring of a plastic filter funnel top, if used, must be rinsed after the funnel has been disconnected. Even with these precautions carry over may occur. To account for potential carry over, the order of filtration must be recorded. Positive samples should then be compared to filtration order. If a positive result of low coliform density is found immediately after one of a higher density, subsequent contamination may have occurred. Repeat samples of both should be taken and a note of explanation of probable subsequent contamination included in any report.

If these recommendations are followed, the probability of obtaining a positive coliform result through subsequent contamination is minimized. None of the recommendations are contrary to present regulations or guidelines and can be instituted with little or no added effort or capital expenditure. The importance of obtaining accurate results of coliform densities is obvious. Employment of these recommendations would help insure that accuracy.

4.3 CHANGES IN COLIFORM DENSITIES IN WATER SAMPLES

The accepted procedures for handling samples from water distribution systems allow for a delay of up to 30 hours between the time of sample collection and when the sample is filtered (Bordner et al. 1978). In this project samples were collected in the morning and filtered in the afternoon. The maximum time between sample collection and filtration was eight hours and most samples were filtered within 2 to 6 hours after collection. However, we did obtain some information on changes in coliform densities in water samples as a result of two types of experiments.

4.3.1. Preservation Techniques

As a part of the development of procedures to be used for this project, before the field sampling was initiated we tried an experiment with sample preservation techniques. The samples for this experiment were obtained from system WD, a large water system which was not included in the field studies. The reason for use of this system was that the laboratory director knew of a location which consistantly gave samples with a low coliform density.

All samples were obtained from a single fire hydrant. One gallon sample bottles were used. The preservation techniques used were 1. chilling the sample bottles in an ice chest (there was no opportunity for immersion of the bottle tops due to the size of the bottles), 2. adding 2.4ml of 10% thiosulfate solution for a 3 liter sample, and 3. adding 3 gm of peptone and 15 gm boric acid for a 3 liter sample. The third treatment was based on the preservation medium of Brodsky et al (1978) as modified by Lariviere (1978). Chilled and unchilled samples were used for each of the preservatives and for samples with no chemical preservatives, giving a total of six different treatments.

Eighteen three liter samples were collected in one gallon bottles. This gave three samples for each treatment. The sample bottles were numbered; bottles 1 through 6 had no preservatives added and 1-3 were held at ambient temperature while 4-6 were chilled, and so forth. The time sequence of filling the numbered bottles was obtained from a table of random numbers. During sample collection the temperature of the water from the hydrant decreased from 18°C to 16°C. The temperature of the water in the bottles held at ambient temperature increased to 27°C over the first 24 hours and held constant for the second 24 hours. The temperature of the water in the chilled samples decreased to 6°C in the first 24 hrs and then held constant.

Three replicate 100ml aliquots were taken from each sample bottle at 3, 6, 21, 27, and 48 hours after sample collection and filtered through membrane filters. The filters were incubated at 35°C for 22 hours and coliform counts were made. The presumptive coliform colonies were not verified in this experiment because our interest was solely in comparing the preservation techniques. At the same time the aliquots for coliform counts were taken, three 1mℓ aliquots were taken for standard plate counts.

The standard plate count results are given in Table 4.9 and the coliform results in Table 4.10. Analysis of variance was used to detect significant differences in the rows and columns of these tables. These differences are exhibited by differences in the 95% confidence intervals for the individual counts which are given in the tables. The standard plate counts are discussed first because those results are useful in understanding the coliform results.

There were variations in the bacterial densities in the water during the 45 minute period that it took to fill the 18 gallon bottles. In spite of the fact that the order of filling the bottles was randomized to try to avoid significant differences among the treatments initially, one treatment (peptone borate preservative held at ambient temperature) had a significantly lower standard plate count to start with. No significant changes in the standard plate counts between 3 and 6 hours after collection were found. The standard plate counts in the samples held at ambient temperatures with no preservative and with thiosulfate increased greatly during the first day after collecton and apparently continued to increase the second day after prevented collection of precise data on this point. The samples with no preservative or with thiosulfate show no significant changes in 48 hours. The peptone-borate preservative caused a decrease during the first day followed by an increase the second day. It may be speculated that borate at the concentration used was toxic to some of the bacteria but other types of bacteria which are a part of the standard plate count can grow in the presence of borate.

The coliform data are more difficult to interpret because the mean counts were much lower and the variability of the counts was such that statistically significant differences are rare. It appears that the coliform density was decreasing in the samples with no preservative and with thiosulfate which were held at ambient temperature. However, since the standard plate counts were increasing greatly for these two treatments, the apparent decrease may have been the result of interference from the non-coliform bacteria growing in those samples.

4.3.2 Second Day Filtration

During the sampling of system LB, a question arose as to why we were finding so many samples with coliforms whereas the twelve samples per week collected by system personnel never showed any coliforms. They allowed their samples to sit at ambient temperature, sometimes in direct sunlight, for up to seven hours while they were being collected and then refrigerated them for about 20 hrs before they were filtered. Our samples were placed in an ice chest within an hour after collection and filtered the same day.

To test coliform die-off in the 24 hours after collection, we collected some 250ml samples and saved the rest of the sample after filtering 100ml the day of collection. The samples were stored in a refrigerator. On the day after the samples were collected, 100ml aliquots from those which had coli-

TABLE 4.9

EFFECT OF PRESERVATION TECHNIQUES ON STANDARD PLATE COUNT

Treatment	Standard Plate Count per ml at Indicated Time (95% Confidence Interval)				
	3 hours	6 hours	21 hours	27 hours	48 hours
No Additive Chilled	53 (46.4 to 59.6)	53 (45.3 to 60.7)	42 (37.4 to 46.6)	42 (21.9 to 62.1)	47 (42.2 to 51.8)
No Additive Ambient	46 (42.5 to 50.5)	47 (44.6 to 49.4)	148 (34.5 to 261.5)	337 (68 to 606)	>1000 -
Thiosulfate Chilled	41 (43.3 to 48.7)	45 (39.9 to 50.1)	44 (30.5 to 59.5)	27 (21.3 to 32.7)	53 (42 to 64)
Thiosulfate Ambient	39 (34.2 to 43.8)	44 (38.2 to 49.8)	435 (345 to 525)	1075 (843 to 1307)	>1000 -
Peptone-Borate Chilled	41 (35 to 47)	25 (11.3 to 38.7)	8 (6.7 to 9.3)	12 (6.3 to 17.7)	22 (5.9 to 38.1)
Peptone-Borate Ambient	28 (23.4 to 32.6)	22 (17.1 to 26.9)	5 (1.3 to 8.7)	3 (1.7 to 4.3)	12 (4.7 to 19.3)
All Chilled Samples	45 (40.9 to 49.1)	41 (32.5 to 49.5)	31 (21 to 41)	27 (17.5 to 36.5)	41 (31.1 to 50.9)
All Ambient Samples	38 (34.8 to 41.2)	38 (31 to 45)	196 (82.1 to 309.9)	396 (113.8 to 678.2)	-
All Samples	42 (39.3 to 44.7)	40 (34.7 to 45.3)	113 (44.2 to 174.8)	201 (55.7 to 346.3)	-

TABLE 4.10

COMPARISON OF PRESERVATION TECHNIQUES

Treatment	Coliform Count per 100 ml at Indicated Time (95% Confidence Interval)				
	3 hours	6 hours	21 hours	27 hours	48 hours
No Additive Chilled	5.33 (2.35 to 8.31)	2.78 (1.35 to 4.21)	2.89 (1.28 to 4.39)	1.78 (0.33 to 3.23)	2.00 (0.55 to 3.45)
No Additive Ambient Temp.	4.67 (3.45 to 5.89)	2.56 (1.22 to 3.90)	1.67 (0.52 to 2.82)	1.33 (0 to 3.61)	0.85 (0 to 1.96)
Thiosulfate Chilled	4.56 (2.92 to 6.20)	2.00 (0.36 to 3.64)	2.44 (0.38 to 4.52)	2.33 (0.74 to 3.92)	3.00 (0 to 7.41)
Thiosulfate Ambient Temp.	2.44 (0.36 to 4.52)	2.44 (0 to 5.38)	0.67 (0.12 to 1.22)	1.22 (0 to 2.61)	0
Peptone-Borate Chilled	3.67 (1.36 to 5.98)	2.00 (0.71 to 3.29)	3.00 (1.50 to 4.50)	1.78 (0.30 to 3.26)	1.17 (0.15 to 2.09)
Peptone-Borate Ambient Temp.	4.78 (3.16 to 6.40)	2.33 (0.30 to 4.36)	1.33 (0.38 to 2.28)	1.00 (0.15 to 1.85)	6.33 (0 to 15.80)
All Chilled Samples	4.52 (3.35 to 5.69)	2.26 (1.54 to 2.98)	2.78 (1.30 to 4.26)	1.96 (0.54 to 3.38)	2.05 (0.61 to 3.49)
All Ambient Samples	3.96 (3.12 to 4.80)	2.44 (1.35 to 3.53)	1.22 (0.34 to 2.10)	1.18 (0.13 to 2.23)	2.39 (0 to 8.89)
All Samples	4.24 (3.86 to 4.92)	2.35 (1.73 to 2.97)	2.00 (1.18 to 2.82)	1.57 (0.73 to 2.41)	2.22 (0 to 5.40)

forms were filtered and the filters were incubated on M-Endo medium. The results from these samples are presented in Table 4.11. All but one of the samples from system LB showed a decrease in coliform count and in most cases the coliform count on samples filtered the day after collection was zero.

Later in the project, the question of how large a density could occur in samples which were counted as too numerous to count (TNTC interpreted as > 80/100ml) arose. The second day filtration experiment was tried with several samples with coliforms and 1ml and 10ml aliquots were filtered from the samples which gave TNTC results on the day of collection. These results are also presented in Table 4.11. Significant numbers of samples from systems WH and BL were examined in this way. Coliform counts increased in some samples and decreased in others. On the average there was a slight increase.

It is clear that coliform counts on samples filtered the day after collection are questionable. It is difficult to predict whether the coliform count will increase or decrease in samples within the first 24 hours after collection. This apparently depends upon water quality factors not yet elucidated.

4.4 OBSERVATIONS ON INTERFERENCE

During the sampling of system CV filtration of one liter samples was tried as a method of detecting coliforms at densities lower than 1 per 100ml. As can be seen from Tables A5 and A6 there was essentially no correspondence between finding coliforms in one liter samples and finding them in 100ml samples. Later in the project we found that coliform density changes abruptly in water flowing from a tap and there is no reason to expect such correspondence. However, at the time, finding coliforms in 100ml samples from three locations sampled on 3/13/79 and not finding them in one liter samples from the same locations raised a question. When coliforms were found in 100ml samples from 2 locations on 3/15/79 without coliforms in the one liter samples, the filters from the one liter samples were inspected carefully. They were found to be covered with a film of particulate matter. The filters from the one liter samples from the two positive locations were carefully transferred from M-Endo medium to Lauryl Tryptose broth. Gas was produced from fermentation of the LTB within 24 hours and the presence of coliforms was confirmed by transfer to and production of gas in BGLB. It was concluded that coliforms were present in those samples and the particulate matter had interfered with the development of recognizable coliform colonies on the filters when they were incubated on M-Endo.

The following week, three one gallon samples were obtained from a location in the system which was known to give turbid water samples. A multiple tube fermentation test did not reveal any coliforms in these samples. The samples were spiked with a strain of Enterobacter cloacae which had been isolated from a previous sample from the CV system. Aliquots of 100ml, 200ml, 500ml, and 1000ml were filtered from each of the three samples. The filters were incubated on M-Endo medium. The coliform counts obtained were proportial to the volumes filtered and the coliform recovery gave the same density as calculated from the coliforms added. Thus, it appeared that the particulate matter in the samples, which formed a film on the filter, did

TABLE 4.11

DELAYED FILTRATION EXPERIMENTS

System	Sampling Date	Sample Code	MF Coliform Count	
			Filtered 3 to 6 Hours After Sampling	Filtered 21 to 24 Hours After Sampling
LB-BB	7/10/79	331-1	1	0
	7/17/79	178	10	5
		258-2	TNTC	TNTC
		334-1	26	18
	7/18/79	328-1	9	0
		329-1	10	0
		331-1	3	0
LB-TP	7/10/79	688-2	6	0
		699-1	1	0
		699-2	71	25
	7/17/79	470-2	1	6
		488-1	2	0
		566-1	2	0
		638-1	1	0
		717-1	6	0
	7/18/79	475-1	49	0
		480-1	3	0
		577-1	27	1
		591-1	26	0
SR	10/31/80	113-1	1	0
	11/14/80	111-3	1	2
MW	10/31/80	22-4	1	7
	11/12/80	23-3	26	33
	11/14/80	51-7	5	28
		52-7	3	2
		55-3	14	26
		61-4	2	0
		61-8	7	6
		72-12	1	0
BG	10/31/80	122A-1	8	0
		123A-1	2	0
		123A-4	8	1
	11/20/81	123B-1	13	3
		123B-2	9	23
		123B-5	1	4
		1646-1	89	117
		1646-2	4	8

TABLE 4.11
(continued)

System	Sampling Date	Sample Code	MF Coliform Count Filtered 3 to 6 Hours After Sampling	 Filtered 21 to 24 Hours Sampling
WH	5/27/81	114-2	6	3
		172-1	1	0
		206-1	12	3
		206-2	5	5
	5/28/81	114-1	6	15
		114-2	4	0
		174-2	TNTC	1600
		216-1	3	3
	6/2/81	206-1	2	9
		195-2	47	66
		185-1	1	1
		185-2	TNTC	330
		106A-2	3	1
		133-2	1	1
		132-1	3	0
		143-2	2	2
	6/4/81	163-1	3	TNTC
		163-2	43	21
		198-1	1	1
		198-1	42	33
		106C-2	5	TNTC
		209-2	TNTC	2700
		212-1	1	5
		212-1	1	6
		214-1	15	2
		215-1	2	4
BL	6/16/81	122-1	5	34
		152-1	2	2
		152-2	2	0
		154-1	3	0
		160-1	78	290
		160-2	47	TNTC
		163-2	1	0
		164-2	1	2
	6/18/81	118-1	3	1
		127C-2	2	4
		146-2	6	18
		146A-1	1	1
		154-1	7	9
		154-2	20	36
		156A-1	2	0
		168-1	1	1
		170-2	1	0

not interfere with the recovery of Enterobacter cloacae freshly grown in a laboratory culture. It is also possible that the particulate matter which caused the turbidity in the samples which were spiked with Enterobacter cloacae was different from the particulate matter which caused the turbidity in the earlier samples.

One liter samples were used during the first two weeks of sampling system WH. Of 13 one liter samples of well water collected on 4/26/79, two produced coliform colonies on membrane filters on M-Endo. The other 11 filters did not produce coliform colonies on M-Endo but did produce gas in LTB after they were transferred. Of four one liter samples collected at the well (2 raw water samples and 2 chlorinated water samples) on 5/1/79 none produced coliform colonies on filters on M-Endo but two of the four filters produced gas in LTB. This was also seen on two filters from liter samples from the distribution system collected on 4/26/79. We were unable to find a way to eliminate this interference from particulate matter in large volume samples and stopped using one liter samples.

Filters from 100ml samples were inspected carefully for the presence of a film of particulate matter throughout the project. Occasionally, filters with noticeable particulate matter and no coliform colonies after incubation on M-Endo were transferred to LTB. None of these filters caused the production of gas in LTB so, we do not have any evidence of particulate matter interference on 100ml samples.

One liter samples were again used during the sampling of system MI. This system uses well water with a very low pH and all the pipes in the system are plastic. The turbidity of the water in the system is very low and none of the liter samples produced a film of particulate matter on the membrane filters. No coliforms were found in any of these one liter samples.

Le Chevalier et al. (1980) presented some evidence that some bacteria which are enumerated by the standard plate count technique are antagonistic to coliforms and may suppress their growth on membrane filters when present in large numbers. They also cited references to three previous investigations which showed interference with coliform detection by some of the standard plate count organisms. We found evidence of a negative correlation between standard plate counts and coliform densities only for system CV and even for that system the correlation was significant at the 5% level (Goshko et al. 1981). Positive correlations significant at the 1% level between standard plate counts and coliforms were found for systems WH, LB, and BL using the Spearman rank correlation test (Goshko et al. 1981). Thus, what we found was some evidence of an increase in the probability of finding coliforms with increasing standard plate counts rather than any evidence of interference. The standard plate count flora is probably different for different water distribution systems, so it is not necessarily expected that the relationship between coliform count and standard plate count would be the same for all systems.

4.5 VERIFICATIONS OF PRESUMPTIVE COLIFORM AND OTHER COLONIES

Tables 6 and 7 in the various appendices give the results of verifications for individual systems by dates and locations. Both lauryl tryptose broth (LTB) and brilliant green bile broth (BGLB) were used for verification. The results from both broths were consistent throughout the project; i.e., all presumptive coliforms which produced gas in LTB also produced gas in BGLB and, of course, those which did not produce gas in LTB did not produce gas in BG.

The verification protocol described in section 3 was followed with a few exceptions. The number of presumptive coliforms found in systems WH and LB sometimes exceeded our capacity to make verifications and fewer than five colonies were picked from filters which had many colonies. During the sampling of system [illegible] several of the attempted transfers from the membrane filters on M-Endo medium resulted in no growth. This occurred occasionally in other systems. No explanation of this phenomenon was found. The occasional misses on transfers could be due to technician fatigue, but the large number of missed transfers on 10/23/80 may have been the result of some material filtered from the water which interacted with a component of the M-Endo medium to eventually inactivate the coliforms.

The presumptive coliform counts were reduced by the fraction of colonies verified on a sample by sample basis to give the verified coliform counts. In the cases when no coliform colony was isolated for verification for an individual sample the presumptive coliform count was reduced by the overall fraction of colonies verified for that sampling day. In the case of the LB samples for 7/18/79 when no colonies were isolated for verification, the overall fraction of colonies which verified for that week was used to adjust the counts. Since all of the verifications attempted for all samples from that system for July 1979 were positive, this amounted to assuming that all of the presumptive coliforms were verified.

Table 4.12 gives the overall fraction of coliforms which verified for the various systems sampled. The percent verified changes from system to system and from one sampling period to another for some systems. It appears that the percent of the presumptive coliforms which verified is greater in warm weather than in cold weather, but this not entirely consistent for all systems.

Table 4.13 gives the verification results broken down according to the API identifications and also according to the system and sampling period. The data from system DT were further broken down into those which were collected from the water distribution system (DTS) and those which were collected at the water treatment plant (DW). The coliforms found in the water treatment plant were mostly from the raw water or the settled water. The data from the other systems include coliforms from samples collected at the water treatment plant as well as those from the distribution system.

Overall the most frequently identified coliform was *Enterobacter cloacae* (271 isolates) followed by *Enterobacter agglomerans* (165), *Klebsiella oxytoca* (96), *Klebsiella pneumoniae* (77), *Citrobacter* (42), and *Escherichia coli* (25). The *Escherichia coli* isolates were from the DT raw water and Systems WH and BL. The *E. coli* from system BL were from samples collected either at the water treatment plant or from the elevated storage tank. The *E. coli* from

TABLE 4.12

VERIFICATION RESULTS ON SYSTEMS SAMPLED

System	Sampling Period	Verifications Attempted	Positive Verifications	Negative Verifications	Percent Verified
CV	Feb. 1979	21	13	8	61.9
	March 1979	26	25	1	96.1
WH	I April-May 1979	39	39	4	90.7
	II May 1979	140	134	6	95.7
	III May-June 1981	131	120	11	91.6
LB-BB	I June 1979	15	13	3	80.0
	II July 1979	38	38	0	100
LB-TP	I June 1979	60	49	11	81.7
	II July 1979	56	56	0	100
DT	Jan.-Feb. 1980				
	Water Plant	143	131	12	91.6
	Dist. System	27	22	5	59.5
BL	I March-April 1980	54	32	22	59.3
	II June 1980	97	83	14	85.6
MI	May-June 1980	45	33	12	73.3
SR	Sept.-Dec. 1980	145	126	19	86.9
MR	Oct.-Dec. 1980	104	102	2	98.1
BG	Sept.-Dec. 1980	101	91	10	90.1

system WH were in samples collected at the treatment plant and in two specific sites in the community. These results tend to indicate that E. coli does not survive in these water distributions very long. However, apparently species of Enterobacter and Klebsiella do survive and probably reproduce in these systems.

A few of the Enterobacter cloacae and a few of the Klebsiella pneumoniae failed to verify as coliforms. These may be considered to be false negative results.

By far the great majority of the presumptive coliforms which failed to verify were Enterobacter agglomerans. On the other hand over half of the Enterobacter agglomerans isolates did verify as coliforms. Are the Enterobac-

TABLE 4.13

SUMMARY OF VERIFICATIONS AND API IDENTIFICATIONS

	CV		WHI		WHII		WHIII		LBI		LBII		DTW		DTS	
	(+	-)	(+	-)	(+	-)	(+	-)	(+	-)	(+	-)	(+	-)	(+	-)
Escherichia coli			2				4						4			
Enterobacter cloacae	11	2	4		50		25		9		8		3		14	
Enterobacter sakazakii	2				2		4									
Enterobacter agglomerans		5	5	3	4	4	85	9		10	8			4		5
Klebsiella pneumoniae	11				10		4		2		4		10	1	3	
Klebsiella oxytoca	15	1	1		48				7		15		2		4	
Klebsiella ozaenae	1															
Citrobacter	3								3		15		5			
Hafnia alvei															1	
Yersinia intermedia		1														
Aeromonas hydrophila																
Serritia liquifaciens								1					4			
I.D. Not conclusive			1			2										
Unidentified			14	1	20		2	1	40	4	44		105	7		
Presumptive Coliforms	>98		>179		>491		>534		>731		>3649		>1047		>130	
Isolated	53		43		140		131		75		94		143		27	
Verified	43		39		134		120		61		94		131		22	
Not Verified	9		4		6		11		14		0		12		5	

TABLE 4.13 (cont'd)

SUMMARY OF VERIFICATIONS AND API IDENTIFICATIONS

	BLI (+	 -)	BLII (+	 -)	MI (+	 -)	SR (+	 -)	MW (+	 -)	BG (+	 -)	Totals (+	 -)
Escherichia coli	9		6										25	0
Enterobacter cloacae	7		21				59	1	27	1	33		271	4
Enterobacter sakazakii													8	0
Enterobacter agglomerans	7		22	1				6	22		21	5	165	71
Klebsiella pneumoniae	7	6			10								77	7
Klebsiella oxytoca	4												96	1
Klebsiella ozaenae													1	0
Citrobacter	1				1		4		10				42	0
Hafnia alvei													1	0
Yersinia intermedia													0	1
Aeromonas hydrophila								1					0	1
Serritia liquifaciens				1									4	1
I.D. Not conclusive			13	10	1			1		1	1			
Unidentified	4	15	21	2	21	12	43		43			5		
Presumptive Coliforms	>248		234		>495		>719		250		>1259			
Isolated	54		97		45		145		104		101			
Verified	32		83		33		126		102		91			
Not Verified	22		14		12		19		2		10			

ter agglomerans which do not verify false negatives? Are the Enterobacter agglomerans which do verify false positives? Is there any reason that some Enterobacter agglomerans should be considered as microbiological contaminants of drinking water while others are not? There seems to be a lack of correspondence between the functional definition of the coliform group and a taxonomic category which requires some further consideration.

Information on the API code number and reactions on various media of the Enterobacter isolates is given in tables 4.14 and 4.15. The API manual suggests dividing Enterobacter agglomerans into four biotypes based on the reactions to indole (IND) and sucrose (SAC). The information in tables 4.14 and 4.15 are divided according to the API suggested biotypes with the IND+ strains in table 4.14. The IMViC series of tests was carried out for all of the IND+ Enterobacter agglomerans and these results are included in table 4.15. It can be seen that the great majority of Enterobacter agglomerans isolates which did verify as coliforms had IMViC type ++-- which is usually associated with E. coli.

The information on the E. coli isolates is given in table 14.16 and a comparison of the test results used to differentiate between E. coli isolates and Enterobacter agglomerans IMViC type ++-- are given in table 14.17. It is clear that there are distinct biochemical differences between E. coli and Enterobacter agglomerans although the test differences are few.

Enterobacter agglomerans strains were previously considered to be members of the genus Erwinia, but were renamed by Ewing and Fife (1972). Most of them are plant saprophites but a few are opportunistic pathogens of man and other animals. Their sanitary significance as indicators of contamination of drinking water is unclear.

Occasionally non-coliform colonies were picked from the membrane filters, incubated on M-Endo medium, and processed for verification and identification. The non-coliform colonies were described as atypical for red colonies with a dark center but no sheen, red and clear. These results are given in the appendicies. A summary of the results are given in Table 4.18.

Enterobacter agglomerans also turns up relatively frequently in Table 4.18. The numbers in the parentheses with a following + are the numbers of the isolates which produced gas in both LTB and BG. Thus even some of the Enterobacter agglerans which did not pruce sheen colonies on M-Endo media would be counted as coliforms by the fermentation tube technique. Other species of Enterobacter also sometimes produce non-sheen colonies on M-Endo and some of those produce gas in LTB and BG.

The coliform group defined by laboratory methology does not conform exactly to any taxonomic group. Members of the genera Escherichia, Enterobacter, Klebsiella, and Citrobacter usually are coliforms but not in every case. Hafnia Aeromonas, Serritia, and Proteus usually are not coliforms but their reactions to the coliform test procedures are sometimes close enough that they can be confused with coliforms.

The coliform group, as defined by the laboratory test procedures has

TABLE 4.14

SUMMARY OF REACTIONS OF ISOLATES IDENTIFIED AS ENTEROBACTER AGGLOMERANS (IND+)

API Biotype Number	API Profile	Colony on m-Endo	Colony on EMB	Lactose Fermentation	IMViC	Number of Isolates
IND+SAC+	1044133	sheen	nucleated	gas	++--	1
	1044173	sheen	sheen	gas	++--	104
		sheen	nucleated	gas	++--	35
		atypical	sheen	no gas	++--	1
		red	clear	no gas	++--	1
	1044733	sheen	nucleated	no gas	+-++	5
	1044773	sheen	nucleated	no gas	+-++	1
	1045733	sheen	nucleated	no gas	+-++	5
	1045773	sheen	nucleated	no gas	+-++	3
	1244773	sheen	nucleated	no gas	+-++	1
	1245763	sheen	sheen	no gas	+-++	1
	1245773	sheen	nucleated	no gas	+-++	2
IND+SAC-	1044153	sheen	sheen	gas	++--	6
		sheen	nucleated	± gas	++--	2
		sheen	clear	± gas	++--	12
		atypical	nucleated	no gas	++--	1
		atypical	clear	no gas	++--	12
	1044553	sheen	nucleated	no gas	++--	2
		sheen	clear	no gas	++--	7

TABLE 4.15

SUMMARY OF REACTIONS OF ISOLATES IDENTIFIED AS ENTEROBACTER AGGLOMERANS (IND-)

API Biotype	API Profile Number	Colony on m-Endo	Colony on EMB	Lactose Fermentation	Number of Isolates
IND-SAC+	1004133	sheen,atypical	sheen,clear	no gas	2
	1004322	atypical	nucleated	no gas	1
	1004333	sheen	nucleated	no gas	2
	1005132	sheen	clear	no gas	1
	1005133	atypical,red	clear	no gas	2
	1005163	atypical	clear	no gas	2
	1005332	sheen,atypical	nucleated	no gas	5
	1005333	sheen	nucleated	no gas	2
	1005372	sheen	nucleated	no gas	4
	1005373	sheen	nucleated	no gas	3
	1005533	atypical	clear	no gas	1
	1005572	sheen	nucleated	no gas	1
	1006133	red	clear	no gas	1
	1007132	red	clear	no gas	1
	1204173	sheen	sheen	no gas	1
	1204773	sheen	nucleated	gas	10
	1205333	sheen	nucleated	no gas	1
	1205773	sheen	nucleated	gas	1
IND-SAC-	1004103	red	clear	no gas	1
	1004153	red	clear	no gas	1
	1005312	atypical	nucleated	no gas	1
	1005713	sheen	nucleated	no gas	1
	1204113	sheen	nucleated	no gas	1
	1205753	sheen	nucleated	no gas	1

TABLE 4.16

SUMMARY OF ISOLATES IDENTIFIED AS *E. COLI*

API Biotype	API Profile Number	Colony on m-Endo	Colony on EMB	Lactose Fermentation	IMViC Type	Number of Isolates
E.coli A-D	5044152	sheen	sheen	gas	++--	1
E.coli LDC-ODC-	3044552	sheen	nucleated	gas	++--	1
E.coli	4144572	sheen	missing data	gas	++--	1
	5044552	sheen	sheen	gas	++--	1
		sheen	nucleated	gas	++--	2
		atypical	sheen	gas	++--	1
	5044572	sheen	sheen	gas	++--	1
	5144152	sheen	nucleated	gas	++--	3
	5144172	sheen	nucleated	gas	++--	5
	5144572	sheen	sheen	gas	++--	5

TABLE 4.17

COMPARISON OF TEST RESULTS BETWEEN *ESCHERICHIA COLI* AND *ENTEROBACTER AGGLOMERAN* STRAINS WHICH ARE IMViC TYPE++--.

	Escherichia coli	*Enterobacter agglomerans*
API 20E Tests		
Similarities		
Nitrate reduction	+	+
Urease	–	–
Motility	±	±
Gelatin liquification	–	–
β-Galactosidase	+	+
Tryptophane deaminase	–	–
Arginine dehydrolase	–	–
Glucose	+	+
Arabinose	+	+
Meliobiose	±	±
Sucrose	±	±
Mannitol	+	+
Rhamnose	±	±
Insitol	±	±
Sorbitol	±	±
Indole	+	+
Citrate	–	–
Differences		
Lysine decarboxylase	±	–
Ornithine decarboxylase	±	–
Amygdalin	–	+
Other Tests		
Malonate utilization	–	±
Tartrate utilization	+	–
KCN Growth	–	±
Pigment formation (25°C)	–	±
Growth at 44.5°C	±	–

been used for monitoring potable water systems for many decades. It has long been known that, although coliform bacteria are present in large numbers in sewage, there are non-sewage sources of some of the coliform organisms. The lack of exact correspondence between the coliform group and varis taxonomic categories has not been found to be a major impediment to the beneficial use of the coliform group as a monitoring tool.

TABLE 4.18

SUMMARY OF API IDENTIFICATION OF NON-COLIFORM COLONIES

	CV	WH	LB	DT	BL	MI	SR	MW	BG
Escherichia coli					1(1+)				
Enterobacter cloacae	1(1+)	12(2+)	4(2+)	1(4)	5(1+)	1	2	6(5+)	1(1+)
Enterobacter aerogenes		1(1+)							
Enterobacter sakazakii		1							
Enterobacter agglomerans	2	17(1+)	5(1+)	5(1+)	8	2	2(1+)		
API Group 2		2							
Klebsiella pneumoniae			1						
Citrobacter			2						
Hafnia alvei				1					
Aeromonas hydrophila				1	1	5			
Proteus mirabilis	3	3							
Acinetobacter	29	18	4		2				
Pseudomonas fluorescens	5	12		1					
Pseudomonas maltophila		1							
Unidentified				1(1+)		14(1+)	5		4(3+)

SECTION 5

MAXIMUM CONTAMINANT LEVELS FOR COLIFORM BACTERIA AND THEIR RELATION TO FREQUENCY DISTRIBUTIONS

5.1 INTRODUCTION

The Maximum Contaminant Levels (MCL) for coliform bacteria in the National Interim Primary Drinking Water Regulations (U.S.E.P.A. 1976) are based on two rules. The rules considered here are those which apply when the MF technique is used. The first, "the average rule," states that the MCL is 1 coliform per 100ml "as the arithmetic mean of all samples examined per month pursuant to § 141.21 (b) or (c)". The second or "maximum rule" states that the MCL is exceeded when more than 4 coliforms per 100ml are found in more than 1 sample when less than 20 samples are examined per month or in more than 5% of the samples when 20 or more samples are taken per month. Thus the MCL's are based on two factors of the frequency distribution of coliform bacteria: the mean and the probability of obtaining counts >4 per 100ml. However these two factors are not independent of one another. Such dependence is obvious in some instances but more subtle in others. In this section we describe the theoretical dependencies of these two factors and the empirical relationships based on our sampling program. Specifically we address: 1) MCL relationships independent of frequency distribution, 2) frequency distributions considered, 3) results of field sampling, 4) relationship between frequency distributions, sample size, and MCL's, 5) replication of samples, and 6) sample volume.

5.2 CHARACTERISTICS OF MCL'S INDEPENDENT OF SPECIFIC FREQUENCY DISTRIBUTIONS.

Either one or both of the two MCL's may be exceeded depending upon conditions in the distribution system and the monitoring program. Table 5.1 presents hypothetical sampling results for each of these possibilities using a sample size of 10. Results recorded as "too numerous to count" (TNTC) are considered to be >80 per 100ml for calculating the means and variances. For the results presented in rows a through e the average rule is exceeded, but the maximum rule is not. The results in rows a through d may be considered as the minimum conditions needed to exceed the average rule, those in row e represent the maximum variance result, and those in row f represent the maximum average count which exceeds the average rule but not the maximum rule. Overall, sampling results which exceed the average rule but not the maximum would generally be characterized by: 1) a mean in the range of 1 to 3 with a variance approximately equal to or less than the mean; 2) a mean >1 due largely to 1 or a few (<5%) very high counts. In the latter case the variance would be large.

TABLE 5.1

EXAMPLES OF VARIOUS VIOLATIONS OF THE MCL'S WHEN SAMPLE SIZE IS 10

Rule Exceeded		No. of 100ml Samples with Indicated Count														Mean	Variance	Variance/Mean
		0	1	2	3	4	5	6	7	8	9	10	11	79	>80*			
Average	a	0	9	1	0	0	0	0	0	0	0	0	0	0	0	1.10	0.32	0.29
	b	3	4	2	1	0	0	0	0	0	0	0	0	0	0	1.10	0.99	0.90
	c	7	0	0	2	0	1	0	0	0	0	0	0	0	0	1.10	1.85	1.68
	d	9	0	0	0	0	0	0	0	0	0	0	1	0	0	1.10	3.48	3.16
	e	0	0	0	0	0	0	0	0	0	0	0	0	1	0	7.90	624.1	79.00
	f	0	0	0	0	0	0	0	0	0	0	0	0	0	1	≥10.70	>24.35	indeterminate
Maximum	g	8	0	0	0	2	0	0	0	0	0	0	0	0	0	0.80	1.69	2.11
	h	6	2	0	0	2	0	0	0	0	0	0	0	0	0	1.00	1.63	1.63
	i	8	0	0	0	1	0	1	0	0	0	0	0	0	0	1.00	2.11	2.11
Both	j	5	3	0	0	2	0	0	0	0	0	0	0	0	0	1.10	1.60	1.45
	k	6	1	1	0	2	0	0	0	0	0	0	0	0	0	1.10	1.66	1.45
	m	8	0	0	0	1	0	0	1	0	0	0	0	0	0	1.10	2.42	2.20
	n	0	0	0	0	0	0	0	0	0	0	0	0	0	10	≥ 80	≥0.00	indeterminate

*Results recorded as TNTC are assumed to be >80 per 100ml.

The results presented in rows g through i are examples in which the maximum rule but not the average rule is exceeded. In all cases most samples contain zero coliforms per 100ml, and two samples contained >4 coliforms per 100ml. With a sample size of 10, the number of possible results which satisfy these criteria is more limited than those which exceed the average rule but not the maximum rule. This may not hold if the number of samples is greater than 20. In general, the sampling results which exceed the maximum rule and not the average rule may be characterized as: 1) when a preponderance of samples will have no coliforms; 2) when less than 20 samples are examined no more than 4 will have >4 coliforms per 100ml; 3) when 20 or more samples are examined more than 5%, but not more than 20% will have >4 per 100ml; 4) when the percentage of samples with >4 per 100ml is large, the counts of most if not all will be equal to or slightly greater than 4 per 100ml; 5) when the percentage is only slightly greater than 5%, the counts per 100ml may be much greater than 4 for at least one or a few samples. If one sample is recorded as 80 per 100ml, and the maximum rule is exceeded, a minimum of 96 samples are required such that the average rule may not be exceeded.

Both rules may be exceeded in a large number of cases. The results presented in rows j through m slightly exceed the rules and row n represents the opposite extreme. The variety of results by which both rules are exceeded is too great to summarize adequately in a brief table.

Some of the results shown in Table 5.1 are quite improbable (e.g., those in rows a, e, and n). A major purpose of the research described in this report has been to determine the type of frequency distribution of coliforms most commonly found in water distribution systems and to relate this to the MCL's and sampling protocol.

5.3 PREVIOUS WORK ON FREQUENCY DISTRIBUTIONS OF BACTERIA.

5.3.1 Distributions Considered

Five frequency distributions are considered for fitting with sampling results. They are 1) Poisson, 2) Poisson plus added zeroes, 3) Lognormal, 4) Delta, and 5) Negative binomial. Previously, the Gamma-Poisson distribution has also been used (Muenz 1978) but it is essentially the same as the negative binomial. The following is a discussion of the attributes of each.

5.3.2 Poisson Distribution

Pipes and Christian (1978) elaborated on Muenz's approach to evaluating MCL's using simpler frequency distributions: the Poisson and the Poisson plus added zeroes. These simulations were extended somewhat during the first year of the project.

The simplest case occurs when the coliform bacteria are randomly dispersed in the distribution system and the counts on some number of samples would fit a Poisson distribution. This may not be a realistic case, but it forms a lower limit on the number of samples needed to have any given probability of detecting coliforms in the system because it is assumed that the variance is equal to the mean.

If it is assumed that a mean density of one coliform organism per 100ml is a realistic and significant goal, then questions about the relationship between the sample mean and the actual mean density in the distribution system become important ones. Figure 5.1 presents plots of the probability of the sample mean being greater than 1 per 100ml versus number of samples per month for various true mean densities in the system ($\overline{x}$ is the sample mean and μ is the actual mean density). For an actual mean density of 1 per 100ml, the probability that the sample mean will exceed 1 per 100ml is low for a small number of samples but increases to almost 50% as the number of samples increases. A 50% probability of the sample mean being greater than 1 per 100ml means that the water distribution system would fail to meet the regulation 6 out of 12 months on the average. Although the water supply would be judged safe some months, the failure rate over a period of time certainly would not be accepted and adjustments would have to be made to provide better treatment. The failure rate at one sample per month is 26.4% or approximately one month out of four. Even this failure rate would probably not be accepted, so in the long run one sample per month would be adequate to detect a mean coliform density of 1 per 100ml if the bacteria were randomly dispersed in the system. From the two upper curves on figure 5.1, it is clear that a mean density of greater than 1 per 100ml would not be accepted even if only a few samples per month are taken.

The other side of the issue is the possibility of judging a water system with an actual mean coliform density of less than 1 per 100ml as unacceptable. From figure 5.1, it can be seen that if the actual mean density is 0.5 per 100ml and only 1 sample per month is taken, the probability of that sample having more than one coliform per 100ml is about 9% which means that such a system would fail to meet the regulation about once a year. As the number of samples per month is increased the chances of judging a water system with an actual mean density of less than 1 per 100ml as unacceptable decreases rapidly.

Figure 5.2 gives the probability of the samples from a water distribution system failing to meet the maximum rule for coliform bacteria randomly dispersed in the system with actual mean densities of 1, 1.5, and 2 per 100ml. The functions are "saw toothed" because of the integral nature of the samples; that is, 2 samples with 4 or more coliforms is more than 5% of 39 samples per month but it takes 3 samples to be more than 5% or 40 samples per month. Comparison of figures 5.2 and 5.3 shows that the average rule is a more stringent criterion than the maximum rule if the coliform bacteria are randomly dispersed in the distribution system.

Another way of considering the question is to ask what is the probability of obtaining some number of negative samples (no coliforms present) if the actual mean density is 1 per 100ml or less. The left most curve on figure 5.3 gives the probability of obtaining no positive samples for various numbers of samples for an actual coliform density of 1 per 100ml if the coliforms are randomly dispersed in the system. For such a condition the chances of obtaining all samples as negative when more than 2 samples are taken per month is slightly less than 5%.

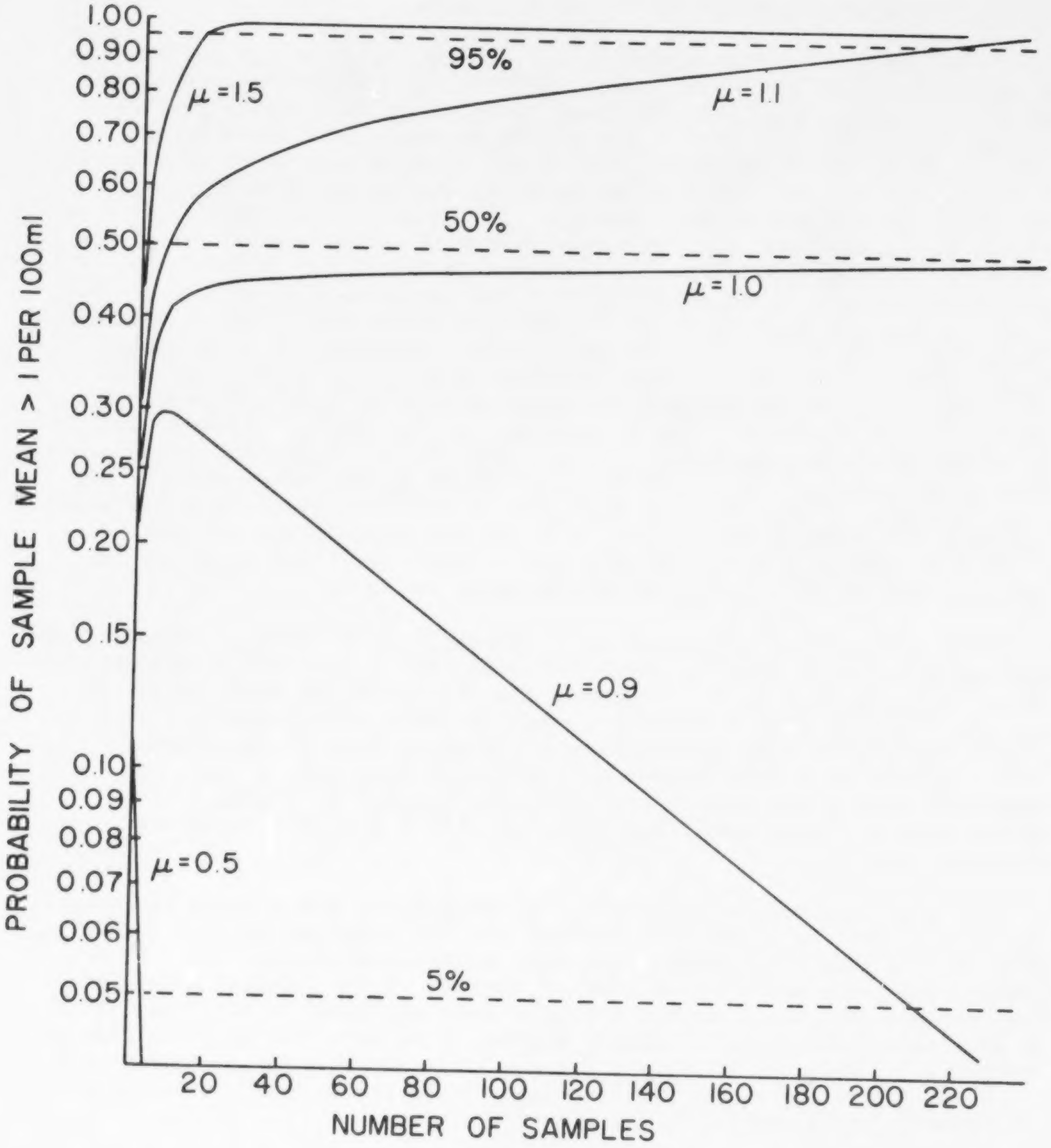


Figure 5.1 Probability of Exceeding Average Rule for Random Dispersion of Coliforms at Different Actual Mean Densities

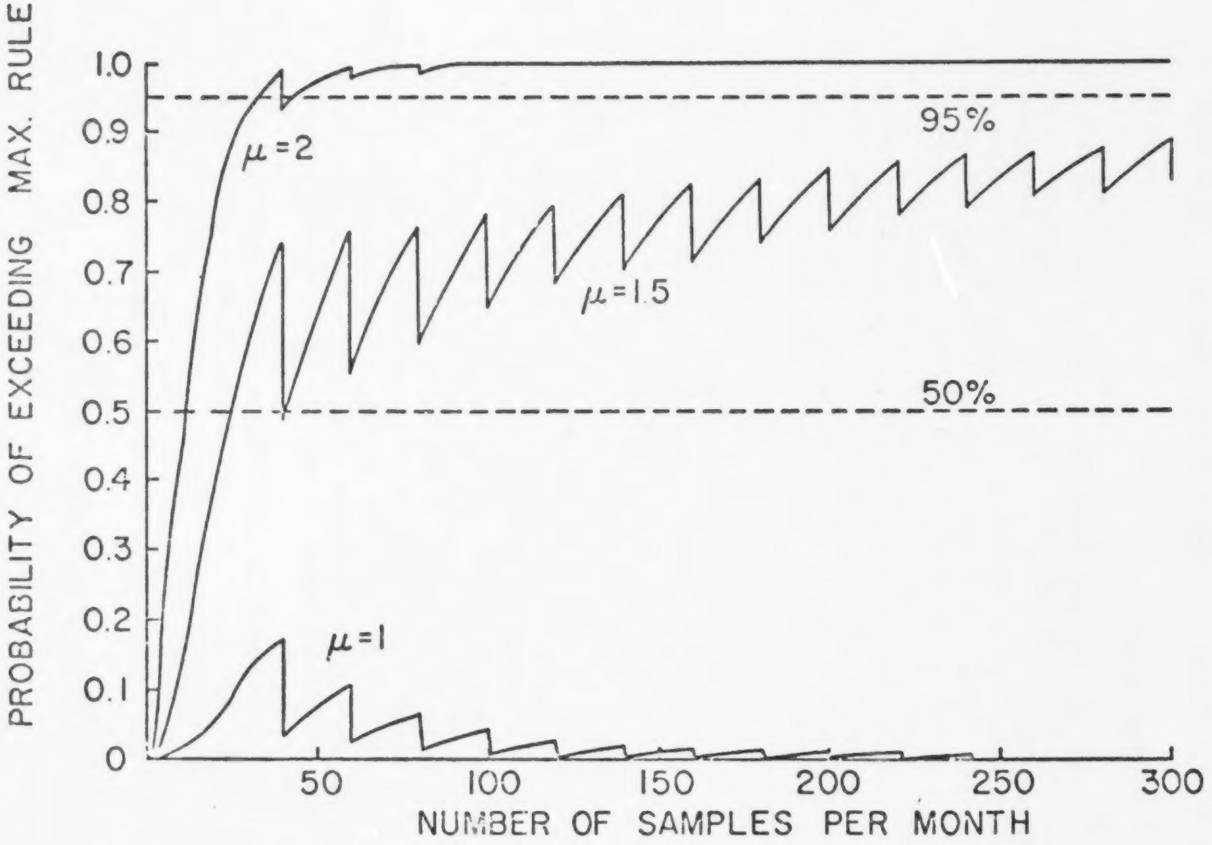


Figure 5.2 Probability of Exceeding Maximum Rule for Random Dispersion of Coliforms in a Distribution System

The overall conclusion from this very simple case is that if coliform bacteria were randomly dispersed in a water distribution system, they would be relatively easy to detect with only a few samples. It would make no difference where the samples were collected and the size of the distribution system would not influence the number of samples required. Unless the actual mean coliform density in the system were considerably less than 1 per 100ml the sample mean would fail to meet the Regulations frequently enough that the operator would be forced to improve the disinfection practices.

5.3.3 Poisson Plus Added Zeroes

The next simplest assumption which can be made about the dispersion of coliform bacteria in water is that some of the water is clean, only a part is contaminated (coliforms present) and the coliforms in the contaminated part are randomly dispersed. The mathematical equivalent of this assumption is the compound binomial-Poisson distribution also called the Poisson plus added zeroes. This is a two parameter distribution. The parameters are the fraction of water contaminated and the mean coliform density of the contaminated fraction. In the earlier paper (Pipes and Christian, 1978) the fraction of water contaminated was represented by $1 - \zeta$ (the Greek letter zeta representing the fraction of water not contaminated) and the mean density of the contaminated fraction was λ. The overall mean density of the water in the system is $\mu = \lambda(1-\zeta)$.

Figure 5.3 gives the probability of obtaining no samples with coliforms versus numbers of samples collected for various values of ζ with $\mu = 1$ per 100ml. To keep $\mu = 1$ per 100ml it is necessary to vary λ with ζ; for instance when $\zeta = 0.5$, $\lambda = 2$ per 100ml, when $\zeta = 0.9$, $\lambda = 10$ per 100ml, etc. The left most curve for $\zeta = 0$ and $\lambda = \mu = 1$ per 100ml is the random dispersion of coliforms throughout the distribution system as previously discussed. From Figure 5.3, it may be seen that with 50% of the water contaminated it is necessary to collect at least 6 samples to have a 95% probability that at least one sample will have coliforms, with 10% of the water contaminated 30 samples are needed, with 5% of the water contaminated 60 samples are needed, and with 1% of the water contaminated 300 samples are needed. From this it is clear that it is never possible to take enough samples to be absolutely sure that there are no coliforms in the system. It is possible to discuss the probability that some small fraction (such as 5% or 1%) of the water is contaminated even for a large number of samples with no coliforms.

This model behaves the same in relation to the average rule as does the random dispersion model for $\mu = 1$, it just requires more and more samples for the probability of the sample mean to approach 50% as ζ gets larger and larger (Pipes and Christian 1978). Also if $\mu > 1$ the probability of the sample mean being greater than 1 approaches 100% for large numbers of samples.

5.3.4 Assumptions Underlying the Simplistic Models

In these simplistic models bacteria are treated as if they are single, inert beads floating around individually in the system. Bacteria are not inert. If a chlorine residual is present, they will die-off in the system. Some are able to multiply in the system. They may occur as clumps of two,

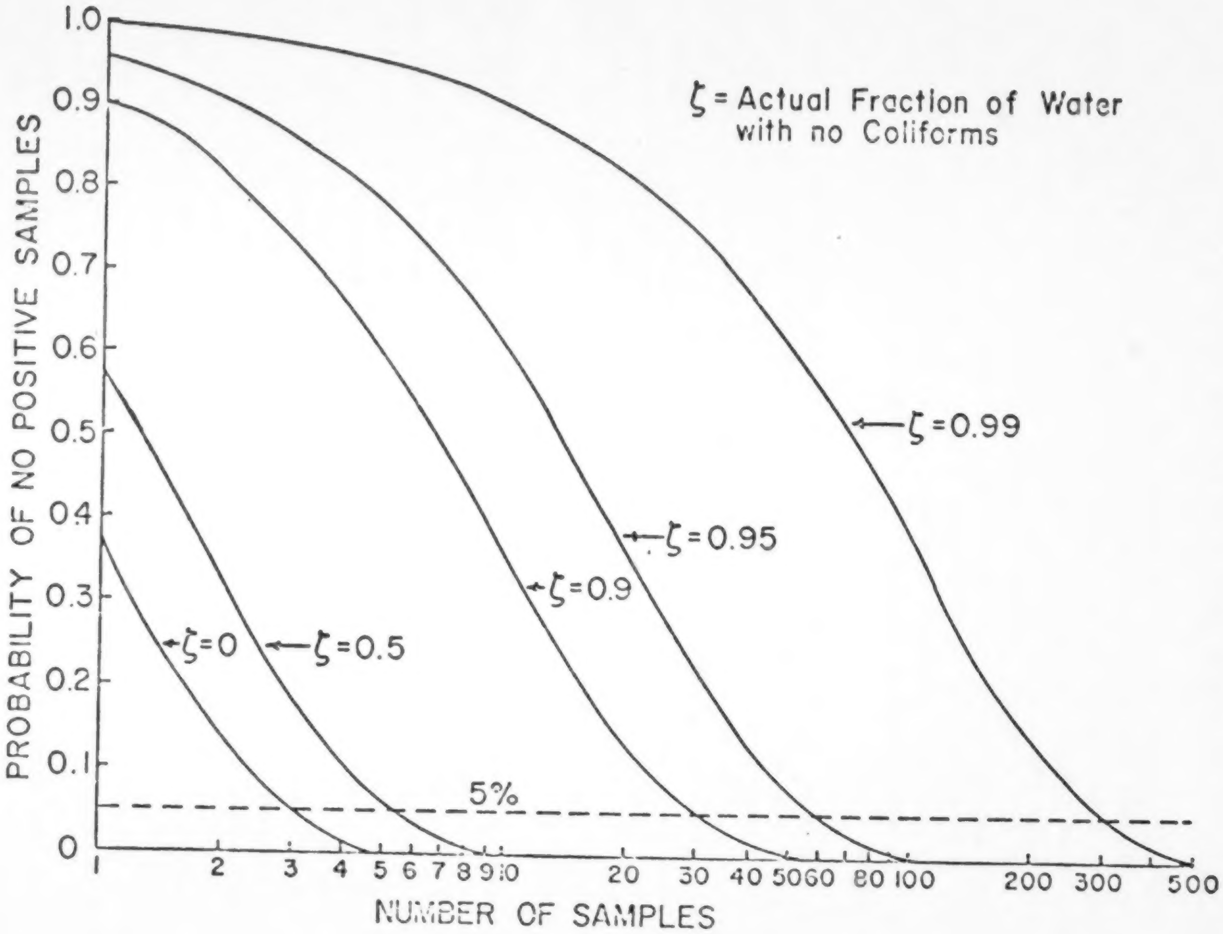


Figure 5.3 Probability of Obtaining No Positive Samples for Actual Mean Density of 1 per 100ml

three, four or more cells adhering together. If the bacteria are growing slowly, it is more likely that the bacteria will occur as clumps rather than as single cells. Also in the presence of chlorine, it is likely that bacteria in clumps would survive longer than individual cells. If the bacteria are in clumps widely dispersed in the distribution system and it is assumed that the shaking of samples required before analysis breaks up the clumps, it would be expected that most of the samples would have no coliforms and that the samples with coliforms would have several.

The second assumption underlying these simplistic models is that one frequency distribution can describe the dispersion of coliform bacteria throughout the system. There is no particular reason to believe, a priori, that this should be true. It has been considered as a hypothesis to be tested. If the hypothesis were confirmed that the coliform counts from an entire system did fit a single random frequency distribution, this would suggest a single mechanism of contamination. However, distribution systems with different pipe materials in different areas, or different ages of pipe or differences in pressure probably are subject to different mechanisms of contamination. A more complex frequency distribution may be observed if this is the case.

The same argument can be made with respect to the period of time over which the samples are collected. Time is not a variable in the simplistic models. No differentiation is made among samples collected over an hour's time or a day or a week or a month. The dispersion of coliform bacteria in a water distribution system may change with time. One of the aims of this project has been to determine the periods of time over which the microbial quality of water in a distribution system is likely to remain constant and attempt to identify significant occurrences in the operation of the distribution system which could cause changes in the level of contamination and the dispersion of the coliforms in the system.

The simplistic models provide some insight into the general relationship between the actual mean density of coliforms in a water distribution system and the sample mean as a function of the number of samples collected. If the actual mean density in the distribution system is 1 per 100ml and enough samples are taken, the sample mean density will be greater than 1 frequently enough that the disinfection practices will have to be improved. "Enough samples" may be very few (3 or 4) if the coliforms are randomly dispersed throughout the system or several hundred if only a small fraction of the water is contaminated. In these simplistic models, the number of samples needed is independent of the size of the distribution system. However, if the microbial quality of the water varies from one part of the system to another, a dependence between the number of samples needed and the size of the system might be shown. A point which has been investigated is how large a portion of a distribution system will show consistent water quality and what features of the system could be used to identify areas of different water quality. It is also clear that the time period over which the microbial quality of the water will remain consistent and the types of changes in the operation of the system that might signal possible changes in microbial quality of the water are important in designing a sampling program.

5.3.5 Lognormal Distribution

The lognormal or variations of the lognormal distribution were investigated during our study. In the lognormal distribution the arithmetic mean is dependent on both the geometric mean and geometric standard deviation. This dependence is exponential rather than linear such that relatively small changes in the geometric standard deviation may produce large changes in the arithmetic mean. This may be seen in Figure 5.4. In this figure the arithmetic mean is plotted within the plane formed by the log geometric mean and log geometric standard deviation. In this case the line designated as "1" is the arithmetic mean density of 1 per 100ml. From the slopes of the lines it can be seen that the geometric standard deviation plays a larger role in establishing the arithmetic mean than does the geometric mean. Thus with a lognormal distribution the variability between samples is important in determining whether the average rule is exceeded. In this distribution samples with large densities are to be expected to occur with a probability not much less than that for low densities. Thus the probability of exceeding the maximum rule increases with an increase in the geometric standard deviation. This is shown in Figure 5.5. While the geometric mean remains constant the probability of obtaining counts >4 increases as the geometric standard deviation increases. The discussion above is theoretical and based on an infinite number of samples. The relationships as functions of number of samples examined are discussed later.

For the lognormal distribution the "average rule" is more stringent than the "maximum rule" when a large number of samples are taken. That is the probability of exceeding the "average rule" is always greater than that of exceeding the "maximum rule." This is shown graphically in Figure 5.6. This may not occur with small sample sizes, as shown later but does indicate redundancy of the rules.

The lognormal distribution assumes no limits as to the upper and lower boundaries of coliform densities. Obviously, however, these limits exist and may be recognized at two levels: first, the actual limits of coliform densities and, second, the sampling limits. The upper limits of coliform densities are the maximum densities which could occur in the water. If coliforms are growing in the water, this could be considered the carrying capacity of the population dependent on the nutrient availability in the water, the growth yield of the bacteria, and physical and time constraints on growth such as temperature and residence time of the water in the system. If coliforms recovered in the water have grown elsewhere (e.g., pipe walls or ground water) the carrying capacity of the growth location and entry mechanism must be taken into account. The actual lower limit is zero for the entire volume of water in the system.

The detection limits depend on the volume of samples. If a standard volume of 100ml is used, the lowest positive count is 1 per 100ml. The upper limit is 80 per 100ml (Bordner et al 1978). To assess the lognormal distribution the data are considered truncated above and below these limits. Samples with 0 per 100ml are considered as <1 per 100ml, and TNTC samples are considered as >80 per 100ml. No distinctions are made as to how much lower or higher than these limits the densities may be. These limits may be

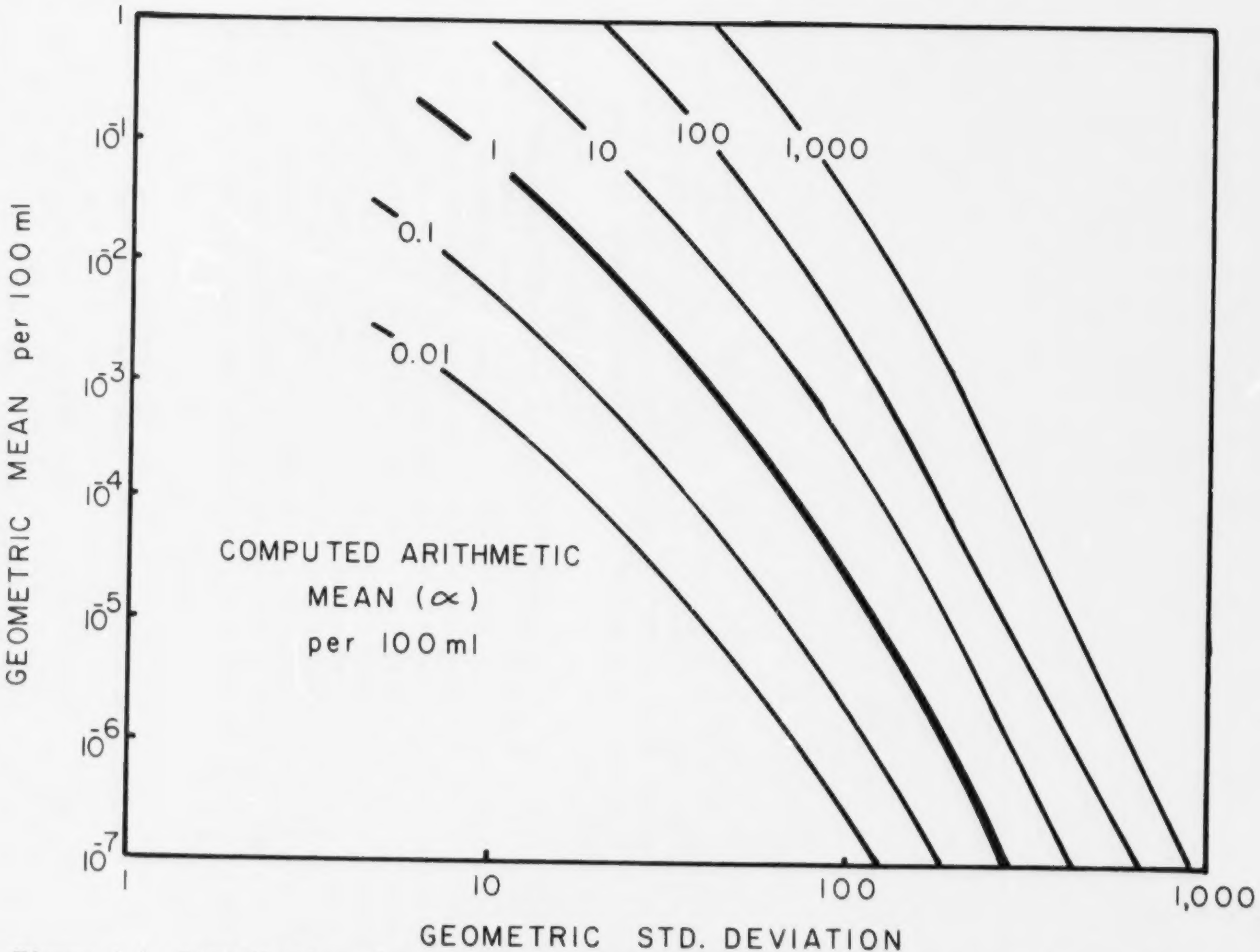


Figure 5.4 Distribution of computed arithmetic mean (μ) from the geometric means and geometric standard deviations in lognormal distributions. Isobars represent different arithmetic means.

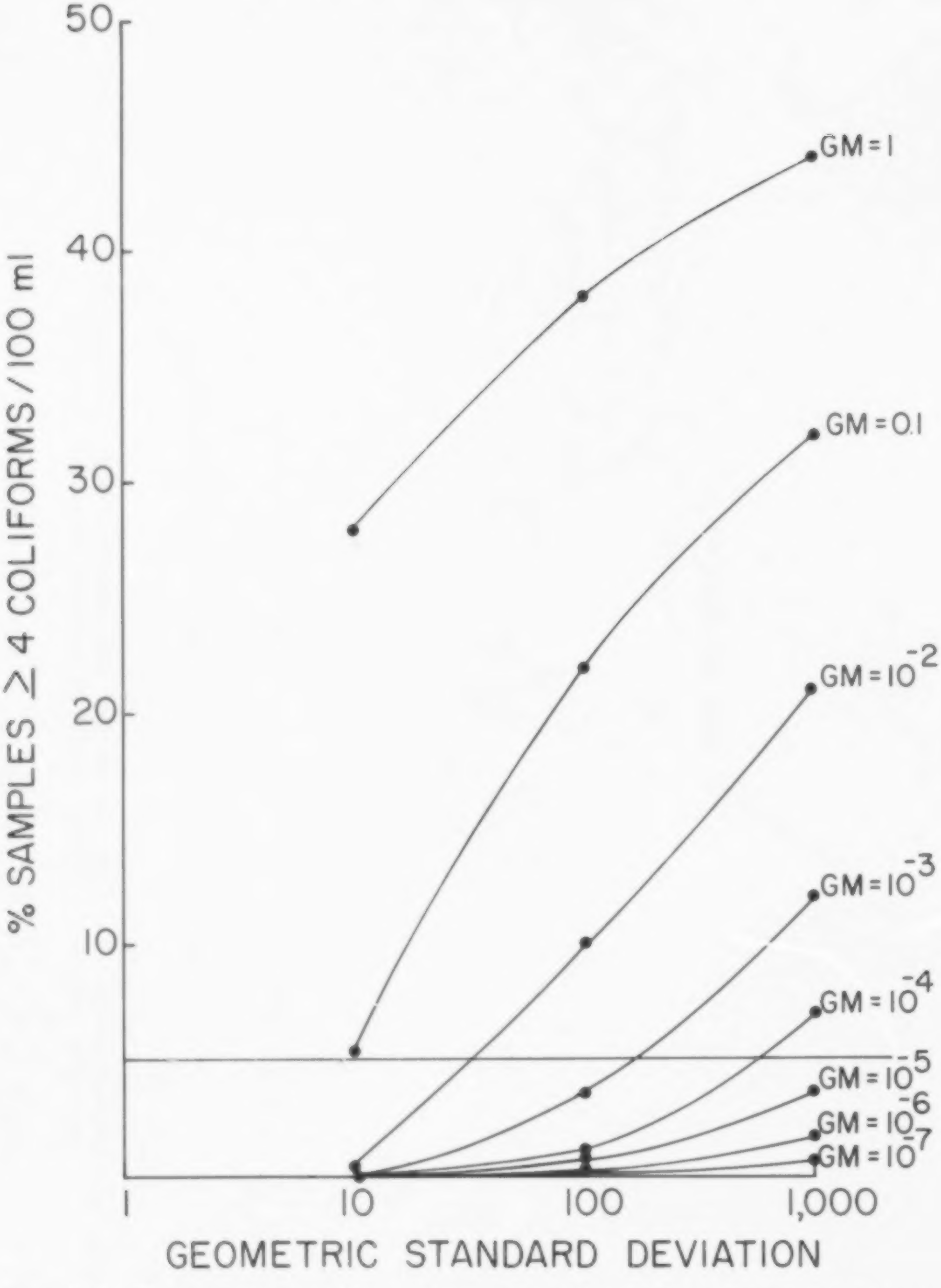


Figure 5.5 The percentage of samples with ≥ 4 coliforms/100ml as a function of the geometric mean (Gm) and geometric standard deviation.

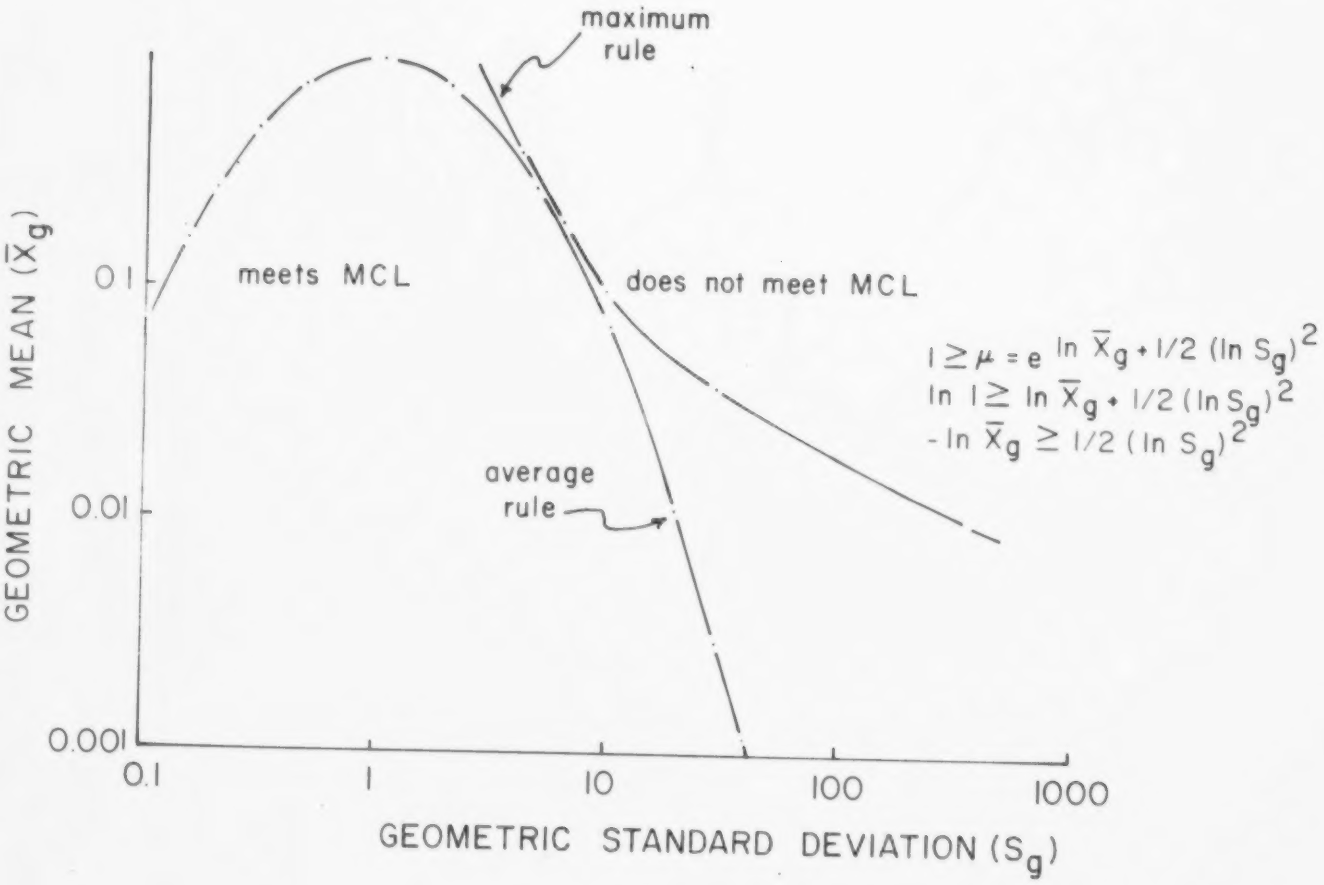


Figure 5.6 Likelihood of exceeding the average and maximum rules for lognormal distributions when a large number of samples are taken per month.

extended by multiple and larger volume sampling as will be discussed later.

5.3.6 Negative Binomial Distribution

Muenz (1978) assumed that the dispersion of coliform bacteria in a water distribution system could be described by a compound gamma-Poisson frequency distribution which is mathematically equivalent to using a negative binomial distribution. Thomas (1952) suggested the use of the gamma distribution for averaging coliform data from water distribution system samples and later (Thomas, 1955) showed that the negative binomial distribution was equivalent. Pipes et al (1977) found that coliform counts from a raw water supply would fit either a negative binomial or a lognormal distribution. Recently, El-Shaarawi et al (1981) used the negative binomial distribution to describe bacteriological data from Lake Erie. The negative binomial distribution is similar to the Poisson distribution but has a variance much greater than the mean. Muenz (1978) interpreted a high variance as indicating greater heterogeneity of the water distribution system with respect to microbial water quality. Part of his results are reproduced in Figure 5.7. In his terms, "probability of acceptance" means the probability of obtaining a sample mean of less than 1 per 100ml for n samples and the "coefficient of variation" is the ratio of the variance to the mean. If the true mean coliform density of the water is 1 per 100ml, the sample mean will be less than or equal to 1 only about half the time if 20 samples per month are collected. However, if fewer samples are collected the chances of getting a sample mean of less than or equal to 1 per 100ml are greater especially if the variance is high.

The negative binomial distribution is commonly used to describe clumped or aggregated patterns of dispersion. The negative binomial may be derived as a compound distribution. It models a random or Poisson distribution of clumps or clusters in which the pattern of coliforms in each clump is logrithmic (Pielou 1969). The negative binomial has three parameters but, since the three parameters are interdependent only 2 are needed to describe the distribution completely. These parameters are the mean, variance, and a measure of aggregation (k). The true mean and variance are estimated well by the simple calculation of arithmetic mean and variance. The measure of aggregation (k) is derived from the mean and variance such that as the degree of aggregation increases, k decreases. As k becomes large the negative binomial distribution merges with the Poisson distribution. Figure 5.8 shows the relationship of the probability of obtaining samples with various densities (0,1,2,3,>4 coliforms/100ml) as the log k is varied for a true mean of 1 per 100ml. When k is small most samples have 0 coliforms and the probability of obtaining counts >4 is comparable to that of obtaining a count of 1. This is similar to what was found for the lognormal distribution. As k increases, the probability of obtaining a positive sample increases, but most positive samples would have a low density (e.g. 1 or 2/100ml). This is similar to the Poisson distribution. Thus, the flexibility of the negative binomial is demonstrated here.

5.4 DATA FITTING

The following frequency distributions were examined: Poisson, Poisson plus added zeroes, lognormal, and negative binomial. The following is a

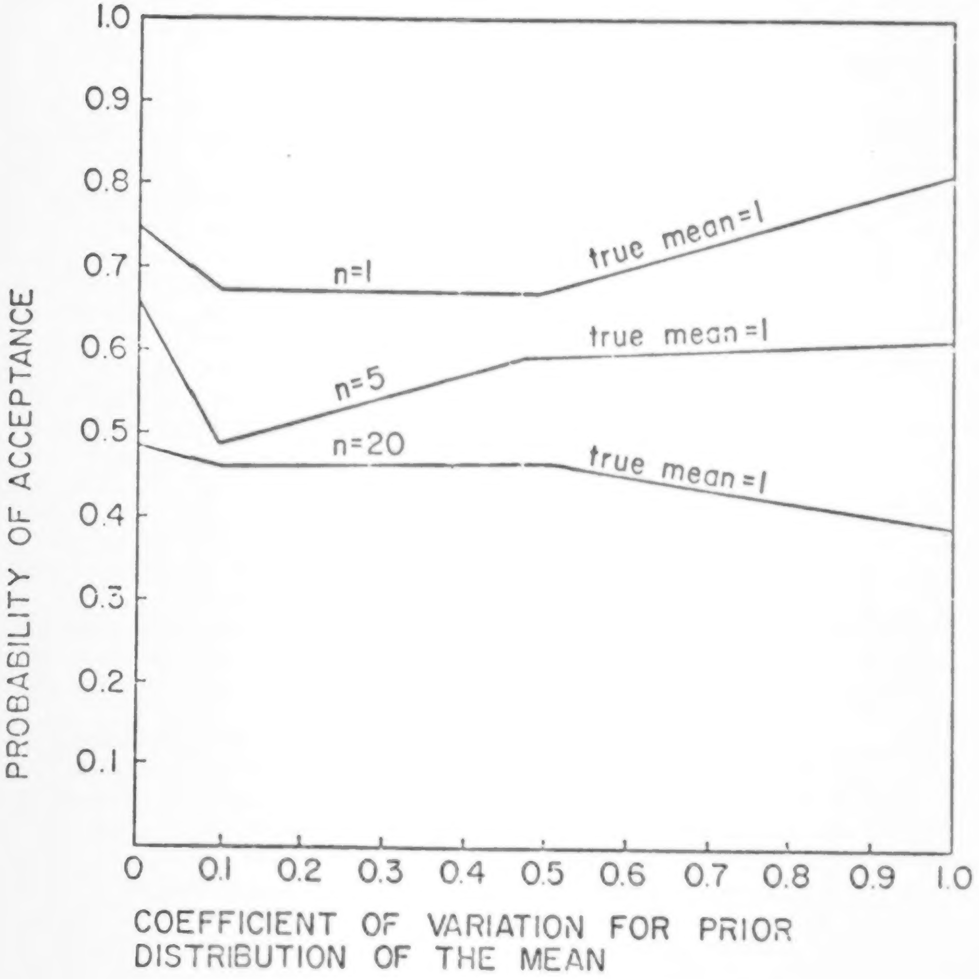


Figure 5.7 The Probability of Obtaining a Sample Mean of 1 for a True Mean Density of 1 as a Function of Number (n) of Samples and Variability of the System. (after Muenz, 1978).

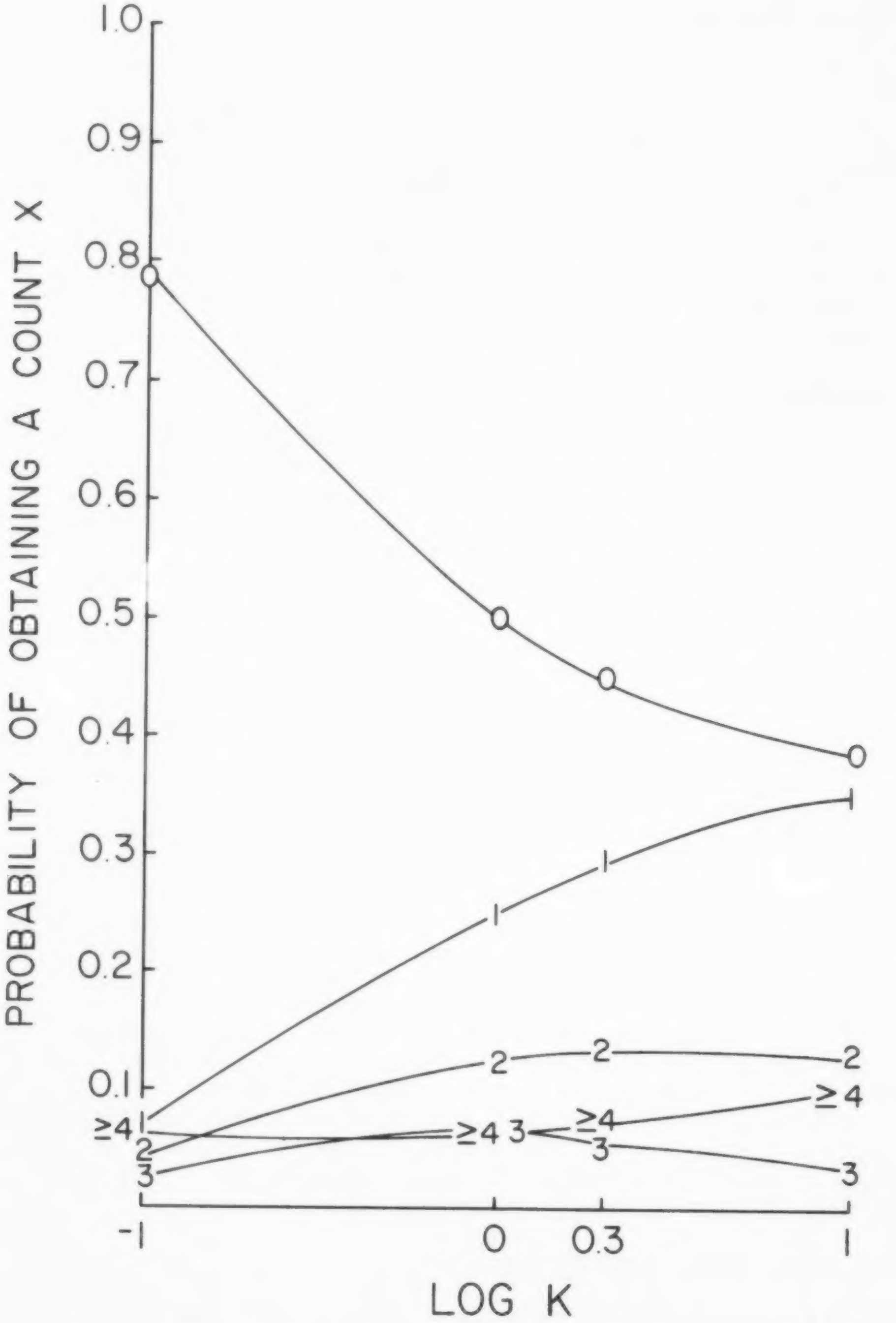


Figure 5.8 Probability of obtaining a particular count (X) when $\mu=1$ and K is varied from 0.1 to 10 in a negative binomial distribution. X=0, 1, 2, 3 $\geq$ 4.

discussion of the applicability of each.

5.4.1 Data Sets

Nine water systems were sampled during the project's three years. Information on sampling each system is summarized in Table 5.2. More information on each is presented in the Appendices. Total coliform results from each system were used to assess the applicability of the various frequency distributions. In some cases the sampling results were divided into shorter periods of time for analysis of time dependent changes in water quality. System LB was divided into two major areas on the basis of separate well water sources. These areas are denoted as Brant Beach (BB) and Terrace (TP).

5.4.2 Poisson Distribution

The simplest assumption as to the frequency distribution of coliforms is that they occur randomly within the system. This is a Poisson distribution in which the mean coliform density is equal to the variance. A test of the applicability of this distribution is the testing of the null hypothesis of equality of mean and variance. This is accomplished by a χ^2 test using Fisher's Index of Dispersion (D^2) where $D^2 = (n-1)S^2/\bar{x}$. The term n is the number of samples, S^2 is the sample variance, and $\bar{x}$ is the sample mean. Another method of testing the hypothesis that the coliforms are randomly distributed is the KolmogorovSmirnov test for goodness of fit (Sokal and Rohlf 1969). In this test the observed cumulative frequency distribution is compared with the expected Poisson cumulative frequency distribution based on the sample mean. The maximum difference between the two curves divided by the number of samples (Dmax) is the test statistic. Each of these procedures were used to test the applicability of the Poisson distribution. The number of samples used for these analyses was less than the total number taken. Only locations in which a constant number of replicates could be analyzed were used. A minimum of duplicate 100ml samples per location was used to restrict unreplicated locations. In systems SR, MW and BG the replication was 2-100ml and 2-200ml samples. Each replicated location was used as a sample for computation.

The results of the analyses for the Poisson distribution are shown in Table 5.3. Mean coliform densities ranged from 0.11 to >9.38 per 100ml including 9 systems in compliance and 6 in violation of the average rule. Both methods of testing failed to accept the hypothesis of a Poisson distribution of confirmed total coliforms in any system. Thus, calculations as to the effectiveness of sampling in determining bacteriological water quality should not be based on the assumption of randomness.

5.4.3 Poisson Plus Added Zeroes Distribution

The second frequency distribution considered assumes that some portion of the water is devoid of coliforms and that in portions of water containing coliforms the bacteria are dispersed randomly. To test this hypothesis the Fisher's Index of Dispersion was used on the means and variances of locations which gave positive coliform results. The results of these analyses are shown in Table 5.4. The fractions of uncontaminated water ranged from 0.585 to

TABLE 5.2

SUMMARY OF WATER SYSTEMS SAMPLED

System	Sampling Divisions	Sampling Dates	No. of Locations with valid Samples	Symbols
Coatesville		2/1-3/15/79	225	CV
Woodbury Heights	I	4/24-5/3/79	66	WH I
	II	5/8-5/17/79	154	WH II
	III	5/22-6/4/81	170	WH III
Long Beach Island	Brant Beach I	6/7-6/21/79	92	BB I
	II	7/10-7/19/79	99	BBII
	Terrace I	6/7-6/21/79	122	TP I
	II	7/10-7/19/79	146	TP II
Downingtown		12/27/79-2/12/80	174	DT
Brooklawn	I	3/4-4/24/80	236	BL I
	II	6/16-6/25/81	169	BL II
Mt. Idy		5/6-6/5/80	207	MI
Spring Run		10/16-12/2/81	144	SR
Marshallton Woods		10/8-12/11/80	99	MW
Bradford Glen		9/25-12/2/80	46	BG

TABLE 5.3

APPLICABILITY OF THE POISSON DISTRIBUTION TO THE FREQUENCY DISTRIBUTION OF CONFIRMED TOTAL COLIFORMS IN WATER DISTRIBUTION SYSTEMS

System	Number of Locations Sampled	Mean Density[a] per 100ml	Variance[a]	D^2[b]	Dmax[c]
CV	225	0.11	0.67	1364	0.956
WH I	66	0.40	3.62	588	0.909
II	154	> 1.52	> 25.21[a]	> 2538	0.677
III	170	> 1.76	> 72.95	> 7004	0.869
BB I	92	0.51	15.41	2750	0.880
II	99	> 9.38	> 439	> 4587	0.789
TP I	122	> 2.58	> 59.06	> 2770	0.766
II	146	> 6.16	> 274.9	> 6471	0.801
DT	174	> 0.30	> 9.40	> 5421	0.954
BL I	238	> 0.35	> 26.89	> 18208	0.967
II	169	> 0.50	> 23.24	> 7809	0.893
MI	207	> 0.40	> 30.92	> 15924	0.986
SR	144	> 0.93	> 25.72	> 3955	0.840
MW	99	0.41	3.41	809	0.818
BG	46	> 2.15	> 69.73	> 1459	0.813

[a] The > symbol has been used when at least one sample was "too numerous to count." The density for such a sample was considered >80 per 100ml.

[b] All systems had a probability of having a Poisson distribution of $p<0.005$ when Fisher's Index of Dispersion (D^2) was analyzed.

[c] All systems had a probability of having a Poisson distribution of $p<0.01$ when the Kolmogorov-Smirnov goodness of fit test (Dmax) was analyzed.

TABLE 5.4

APPLICABILITY OF THE POISSON PLUS ADDED ZEROES DISTRIBUTION TO THE FREQUENCY DISTRIBUTION OF CONFIRMED TOTAL COLIFORMS IN WATER DISTRIBUTION SYSTEMS

System	Number of Locations (Number with Coliforms)	Fraction of Locations Not Contaminated	Mean Density[a] per 100ml of Contaminated Water	Variance of Contaminated Water	D^2[b]
CV	225(10)	0.956	2.52	10.03	40
WH I	66(6)	0.909	4.42	25.84	29
II	154(61)	0.604	> 3.84[a]	> 55.26	> 863
III	170(26)	0.847	> 11.49	> 337.05	> 820
BB I	92(11)	0.880	4.27	122.52	287
II	99(41)	0.585	> 22.65	> 767.52	> 1355
TP I	122(41)	0.664	> 7.67	> 138.62	> 723
II	146(53)	0.637	> 16.99	> 579.42	> 1773
DT	174(8)	0.954	> 6.56	> 185.39	> 198
BL I	236(8)	0.966	> 10.44	> 790.03	> 530
II	169(18)	0.893	> 4.67	> 209.03	> 761
MI	207(7)	0.966	> 11.89	> 902.08	> 455
MI	207(7)	0.966	> 11.89	> 902.08	> 455
SR	144(36)	0.750	> 3.73	> 94.35	> 885
MW	99(19)	0.808	2.15	14.59	122
BG	46(10)	0.783	> 9.90	> 263.43	239

---

[a] The > symbol has been used when at least one sample was "too numerous to count." The density for such a sample was considered >80 per 100ml.

[b] All systems had a probability of having a Poisson plus added zeroes distribution of $p<0.005$ when Fisher's Index of Dispersion (D^2) was analyzed.

0.966. None of the frequencies of distribution of coliforms could be satisfied by the Poisson plus added zeros distribution. The added assumption of a portion of water devoid of coliforms was not sufficient to explain the dispersion of bacteria. The coliforms in contaminated water were not randomly dispersed. The fact that the D^2's in these analysis were larger than could be explained by randomness indicates that the bacteria are aggregated in some fashion. The probability of obtaining high counts (>4 per 100ml) in samples is greater than would be predicted by the two distributions invoking randomness. Again the calculations based on a simple model are inadequate to describe the effectiveness of sampling.

5.4.4 Lognormal Distribution

The lognormal distribution is useful in fitting data which are aggregated and which have samples containing both small and large densities. This frequency distribution has been used previously for the analysis of coliform data (Velz 1951, Pipes et al 1977). The graphical approach using log-probability paper provides a quick method of checking for a fit. Aitchison and Brown (1957) discuss the rationale underlying this method in some detail. Figure 5.9 gives two log-probability plots of the Coatesville coliform data based on individual samples rather than locations. There are two approaches to analysis of these data. One is to assume that some fraction of the water is not contaminated with coliform bacteria (in this case 0.9781 of the water) and fit the positive counts to a lognormal distribution. This is tantamount to assuming a lognormal plus added zeroes or delta distribution. This approach is represented by the line labeled "percent of 10 positive samples." The mean log for the 10 positive counts is 0.457 giving a geometric mean of 2.86 per 100ml and the log of the standard deviation is 0.522 giving a geometric standard deviation of 3.32. The leftmost line of Figure 5.9 passes through the geometric mean at the 50 percent point and has a slope corresponding to the geometric standard deviation of 3.32. This approach requires three parameters to describe the frequency distribution of coliform counts: namely, the fraction of the water contaminated, the geometric mean, and the geometric standard deviation for the contaminated fraction.

The other approach is to assume that all of the water in the distribution system is contaminated and the coliform counts fit a lognormal distribution, but counts of less than 1 per 100ml are not observed. This approach is represented by the line on Figure 5.9 labeled percent of all 457 samples. This line is replotted on a different scale on Figure 5.10 so that it can be extrapolated to the 50th percentile which represents the geometric mean (3.8×10^{-6} in this case). For this type of plot the ratio of the 84.1 percentile value to the 50 percentile value is the geometric standard deviation (in this case $1.45 \times 10^{-3}/3.8 \times 10^{-6} = 382$). This approach describes the frequency distribution of coliform counts using only two parameters.

The lognormal distribution represents the frequency distribution of a continuous variable. Bacterial counts are discrete rather than continuous, but bacterial density is actually a continuous variable. A bacterium may occur in 150ml of water, or in 386ml of water, or in 6.84 liters of water, etc. These occurrences would give bacterial densities of 0.667, 0.259, and

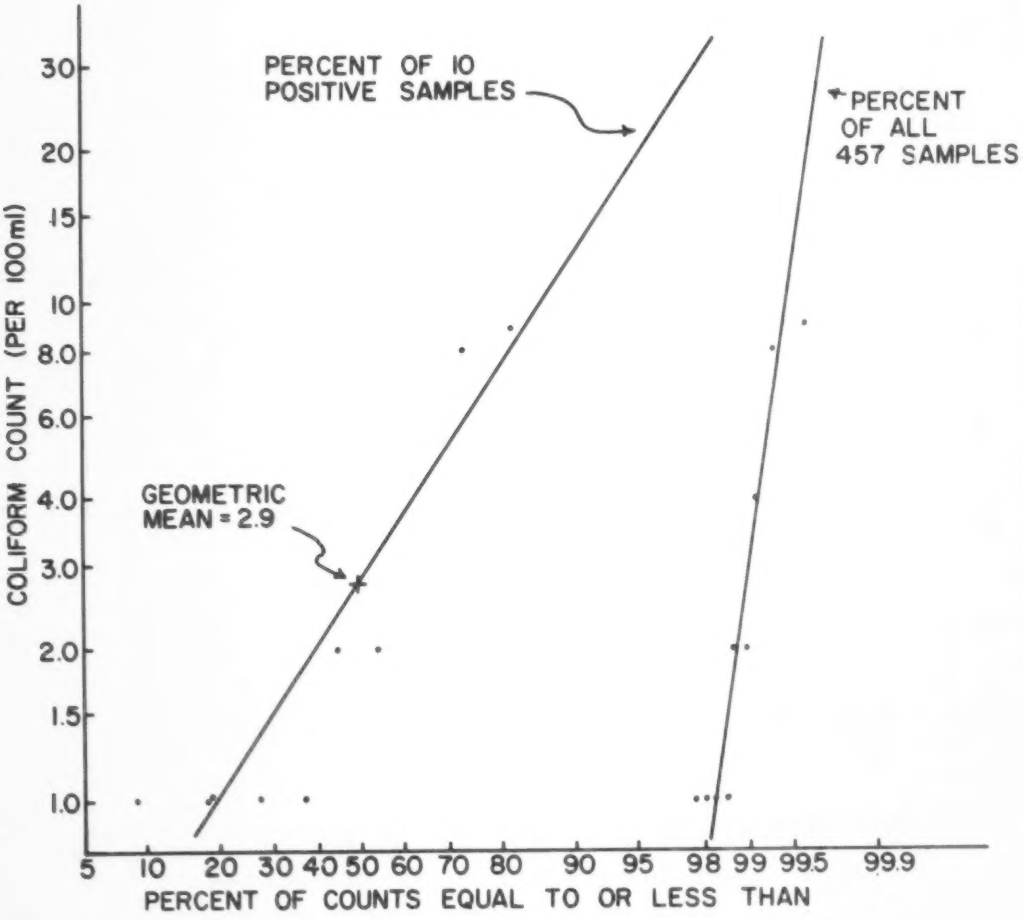


Figure 5.9 Log-Probability Plots of Coatesville Coliform Data

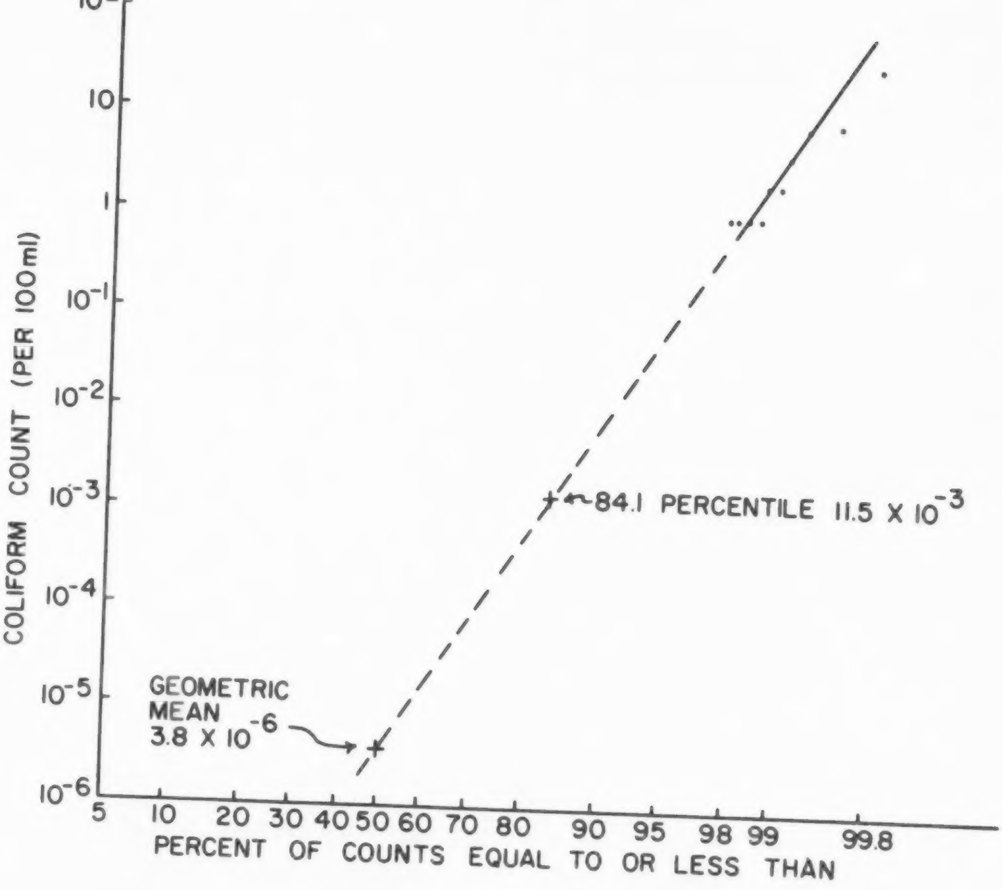


Figure 5.10 Extrapolation of Log-Probability Plot of Coatesville Data for Fitting a Truncated Lognormal Distribution.

0.0147 per 100ml respectively. If a 100ml sample happens to include a coliform, a count of 1 per 100ml is produced, but there is no information about how much more water from that sampling location would need to be examined before another coliform would be found. The practical realities of the sampling and sample processing procedures have imposed an artificial discreteness on the data. Densities of less than 1 per 100ml are not observed but they almost certainly exist. When certain values of a lognormal variable cannot be observed the frequency distribution is said to be "truncated" (Aitchison and Brown, 1957). Truncation may also occur at densities above the range studied (i.e., >80 per 100ml).

It can be seen from Figures 5.9 and 5.10 that either approach gives a reasonable representation of the data. Under most circumstances the three parameter approach gives a better fit, but the two parameter approach would be preferred if it gives a reasonable fit because fewer data are needed to estimate the parameters. For this problem there is an additional implication. If it is assumed that some fraction of the water is not contaminated, this implies that in certain parts of the system taking larger volume samples or taking more samples would not increase the number of coliforms found. If, on the other hand, it is assumed that all the water is contaminated but the variability of the counts gives some 100ml samples with no coliforms, this implies that taking larger volume samples would give a greater probability of finding coliforms and of getting higher coliform counts (if the problem of interference could be overcome).

While the graphic approaches described here are useful in providing a quick estimate of parameters, they are subjective. They rely on an "eyeball" fit of the data by the investigator for the construction of the best fit line and hence the geometric mean and geometric standard deviation. Also, no estimation of variability around the line is easily obtained. A more objective fit may be obtained by regression analysis. This is achieved by the following steps:

1) The coliform densities are ordered from least to most. As multiple samples were taken per site and these replicates may not be considered as independent of one another, we will consider the replicate samples from a site to give one density measurement. To determine density per site, the total volume filtered in 100ml units is divided into the total number of coliforms for each replicate. The mean of the replicates is then determined. Therefore all densities are given as coliforms per 100ml. This technique can provide fractional densities of coliforms. For example, if 2 replicate 100ml sample have counts of 0 and 3, the total density is 3 per 200ml or 1.50 per 100ml.

   Sites with no coliforms are not considered zero but rather some density below the minimum observable density. Following our previous example 1 per 200ml or 0.50 per 100ml would be the minimum observable density. Thus, sites with no coliforms are considered to have a density of <0.50 per 100ml. Truncation is said to occur at 0.50 per 100ml. All sites considered must have the same volume filtered.

If a replicate sample from a site contains ≥80 per 100ml it is too numerous to count. Truncation must occur at this high density because of the uncertainty of the true density of any such replicates. The density at which such truncation occurs is dependent on the data set. The lowest density in our example at which this can occur is ≥80 per 200ml or ≥40 per 100ml. However, a TNTC is often associated with other positive counts (e.g., 5, >80). This would make truncation at ≥85 per 200ml or ≥42.50 per 100ml.

2) Once the densities per site have been ordered, the percent of samples ≤ each density is computed based on n + 1 as 100%. An example of this calculation is shown in Table 5.5. When a density between truncations is represented more than once the percent is computed for each (not shown).

TABLE 5.5

EXAMPLE OF COMPUTATION OF FACTORS TO BE USED IN THE ASSESSMENT OF THE LOGNORMAL DISTRIBUTION BASED ON DATA FROM WH I

Number of Sites	Density per 100ml	% ≤ density	σ of %	ln density
60	<0.5	90.0	1.35	<-0.69
2	0.5	92.5	1.37	-0.69
1	1.0	94.0	1.55	0.00
1	4.5	95.5	1.64	1.50
1	6.5	97.0	1.88	1.87
1	13.5	98.5	2.13	2.60

66 locations sampled; n + 1 = 67
Example computation for % ≤ 4.5/100 ml: (64/67) x 100 = 95.59%.

3) The percent ≤ each density is converted to standard deviation (σ) units to linearize the coordinate. Fifty percent is 0σ and each percent above and below that can be represented by some value in standard deviations. For example, 1 standard deviation is 84.1% total or 34.1% above 50%. The σ units are shown in Table 5.5 for the WH I example.

4) The densities per 100ml are transformed to ln or log densities,

again to linearize the coordinate.

5) Least squares regression is performed on the σ values between the lower and upper truncation densities. The antilog of the intercept (50%) is the geometric mean (GM), and the antilog of the slope is the geometric standard deviation (GSD). A correlation coefficient may be derived indicating goodness of fit, and variances and confidence limits about the GM and GSD may be computed.

In Table 5.6 we show the parameters of the lognormal distribution for WH I, II, and III as calculated by the graphical approach using all locations having equivalent volumes filtered. The geometric means (GM) were lower than the sample arithmetic means ($\bar{x}$) in every case. Arithmetic means (α) were calculated from the lognormal distribution parameters by the equation $\ln \alpha = (\ln GM + 1/2\,(\ln GSD)^2)$. In all α exceeded the sample means. The $\alpha:\bar{x}$ was greatest when the geometric standard deviation was greatest. The lognormal distribution predicted that a small but finite portion of the water contained coliform densities in excess of 80 per 100ml (TNTC). These high densities are recorded as 80 per 100ml in sampling; thus the sample mean is biased low when TNTC's are found.

TABLE 5.6

PARAMETERS OF THE LOGNORMAL DISTRIBUTION OF COLIFORMS IN WOODBURY HEIGHTS, N.J. AS COMPUTED BY THE GRAPHICAL APPROACH USING ALL LOCATIONS HAVING EQUIVALENT VOLUMES FILTERED

Time	Geometric Mean per 100ml	Geometric Standard Deviation	Sample Mean per 100ml	Computed[a] Arithmetic Mean (α)	$\frac{\alpha}{\bar{x}}$
I	1.3×10^{-3}	82	0.40	21.42	53.5
II	0.17	12	$>$ 1.52	3.73	3.5
III	5.5×10^{-4}	336	$>$ 1.76	1.2×10^{4}	6818

[a] $\alpha = \exp[\ln \text{Geometric Mean} + 1/2\,(\ln \text{Geometric Standard Deviation})^2]$

The lognormal distribution parameters fitted by least squares regression for all of the systems are shown in Table 5.7. The applicability of this distribution can be assessed through the coefficient of determination (R^2) and the probability of randomness obtained from the F value of the regression analysis. Only the relationship of BL I was not compatible with the lognormal. Seven positive values used for this regression were 0.5 per 100ml.

TABLE 5.7

PARAMETERS OF THE LOGNORMAL DISTRIBUTION OF COLIFORMS IN WATER DISTRIBUTION SYSTEMS AS COMPUTED BY LEAST SQUARES REGRESSION

System	Geometric Mean (± Standard Error)	Geometric Standard Deviation (± 1 Standard Error)	Coefficient of Determination	F	Probability of randomness (p)
CV	$1.9\text{x}10^{-3}$ $(1.1\text{x}10^{-3}-3.2\text{x}10^{-3})$	29 (22-38)	0.951	156.4	<0.0001
WH I	$1.3\text{x}10^{-3}$ $(3.8\text{x}10^{-4}-4.3\text{x}10^{-3})$	88 (43-181)	0.906	38.7	0.0034
II	0.20 (0.19-0.21)	10 (9.6-10.7)	0.974	1878.9	<0.0001
III	$3.5\text{x}10^{-3}$ $(2.0\text{x}10^{-3}-6.3\text{x}10^{-3})$	92 (63-134)	0.899	142.4	<0.0001
BB I	$3.7\text{x}10^{-3}$ $(1.4\text{x}10^{-3}-9.5\text{x}10^{-3})$	39 (21-71)	0.803	36.6	0.0002
II	0.21 (0.18-0.23)	56 (48-65)	0.960	762.1	<0.0001
TP I	0.15 (0.13-0.16)	19 (17-21)	0.948	687.2	<0.0001
II	$8.4\text{x}10^{-2}$ $(7.4\text{x}10^{-2}-9.5\text{x}10^{-2})$	67 (59-77)	0.963	1059.7	<0.0001
DT	$6.9\text{x}10^{-4}$ $(2.6\text{x}10^{-4}-1.9\text{x}10^{-3})$	53 (31-143)	0.919	57.0	0.0006
BL I	0.50 (-)	0.07 (-)	0.000	$>10^{5}$	<0.0001
II	$8.1\text{x}10^{-3}$ $(4.4\text{x}10^{-3}-1.5\text{x}10^{-2})$	20 (14-29)	0.822	69.4	<0.0001
MI	$1.4\text{x}10^{-6}$ $(9.7\text{x}10^{-7}-1.9\text{x}10^{-6})$	441 (375-518)	0.999	1438.1	0.0168
SR	$9.1\text{x}10^{-3}$ $(7.8\text{x}10^{-3}-1.1\text{x}10^{-2})$	35 (31-40)	0.960	747.5	<0.0001
MW	$8.7\text{x}10^{-3}$ $(5.9\text{x}10^{-3}-1.3\text{x}10^{-2})$	35 (27-51)	0.934	182.5	<0.0001
BG	$1.2\text{x}10^{-2}$ $(6.4\text{x}10^{-3}-2.1\text{x}10^{-2})$	97 (59-160)	0.923	83.4	<0.0001

This would not occur randomly, but did not give a slope reasonable for the lognormal distribution. One TNTC location was obtained from this system and one location had 6 presumptive coliforms per 100ml, but these did not confirm. The regression analysis does not include these samples. If they had been included a lognormal distribution would be reasonable. With the exception of BL I all coefficients of determination were greater than 80%, and all but three greater than 90%. Thus a large portion of the variance associated with the ln density of coliforms was explained by the sigma unit (and hence % $<$ density) coordinate. Based on these analyses, it is quite reasonable to assume that the dispersion of coliforms in water distribution systems is often compatible with the lognormal distribution.

The geometric means of confirmed total coliforms for all systems except BL I ranged from 1.4×10^{-6} per 100ml for MI to 0.21 per 100ml for BB II. The variation around the geometric mean was computed as ± 1 standard error. In the case of WH II the total range of this variation (0.19 - 0.21) was 10% of the mean. However, in other cases the variation exceeded 100% (e.g., DT and MI). As the number of positive samples increased, the variation around the geometric mean decreased.

The geometric standard deviations of confirmed total coliforms for all systems except BL I ranged from 10 per 100ml for WH II to 441 per 100ml for MI. Again an estimate of variation is given as ± 1 standard error. The percent of the range of the standard deviations were from 11 for WH II to 211 for DT. Again as the number of positive samples used in the analysis increased, the variation as a percentage decreased. Thus the precision of the geometric means and geometric standard deviations increases as the number of positive samples obtained increases.

The arithmetic means (α) were computed for the various systems from the parameters derived by regression analyses. These are shown in Table 5.8. They ranged from 0.55 to 693 coliforms per 100ml. In all case α exceeded the sample arithmetic mean ($\bar{x}$). The $\alpha:\bar{x}$ ranged from <1.44 to <394. Thus the sample arithmetic mean appears to underestimate the computed arithmetic mean.

It has been stated that the graphical procedure is useful in obtaining estimates of parameters of the lognormal distribution. If the regression analysis is a more rigorous form of estimation, a comparison of the procedures is important. The geometric means and standard deviations for the 3 WH sample periods may be compared using Tables 5.6 and 5.7. Table 5.8 compares the computed arithmetic means of each. The values of GM's and GSD's estimated graphically may deviate from those estimated by regression. When large deviations occurred the graphical procedure tended to give lower GM's and larger GSD's than the regression procedure. This is because of the tendency to place subjectively large weight on large densities when best fitting a line by eye. When many samples are positive and more evenly dispersed along the line the discrepancies are minimized.

The computed arithmetic means are more dependent on the GSD than the GM. The α varies by $1/2\ (\ln \mathrm{GSD})^2$ but only linearly by the GM. Therefore lack of precision in the estimate of the GSD will more greatly affect α. This is seen in the large α g's for WH I, and WH III in Table 5.8.

TABLE 5.8

RELATIONSHIP OF SAMPLE ARITHMETIC MEAN ($\bar{x}$) AND COMPUTED ARITHMETIC MEANS FROM THE REGRESSION ANALYSIS (αr) AND GRAPHICAL ANALYSIS (αg) FOR THE LOGNORMAL DISTRIBUTION

System	$\bar{x}$	αr	$\frac{\alpha r}{\bar{x}}$	αg
CV	0.11	0.55	5.01	
WH I	0.40	29.31	73.27	21.42
II	$\geq$1.52	2.83	$\leq$ 1.86	3.73
III	$\geq$1.76	96.39	$\leq$54.75	$1.2\text{x}10^4$
BB I	0.51	3.04	5.96	
II	$\geq$9.38	693.03	$\leq$73.88	
TP I	$\geq$2.58	11.45	$\leq$ 4.44	
II	$\geq$6.16	79.85	$\leq$94.13	
DT	0.30	1.83	6.09	
BL I	$\geq$0.35	0.50	$\leq$1.43	
II	$\geq$0.50	0.72	$\leq$1.44	
MI	$\geq$0.40	157.46	$\leq$393.66	
SR	$\geq$0.93	5.06	$\leq$5.44	
MW	0.41	4.83	11.79	
BG	$\geq$2.15	420.37	$\leq$195.52	

The following conclusions may be drawn from the evaluation of the lognormal distribution.

1) Most water distribution systems studied possessed a dispersion of confirmed total coliforms compatible with the lognormal distribution.

2) If the lognormal distribution is assumed, the probability of obtaining samples with densities >80 per 100 ml is appreciable even if most samples show no growth of coliforms. All but 3 systems studied had TNTC samples.

3) Estimation of parameters of the lognormal distribution by regression analysis is superior to the graphical technique.

4) The sample arithmetic mean may deviate considerably from the arithmetic mean computed from the parameters of the lognormal distribution.

5) Precision of parameter estimates is very important to estimating the geometric mean or the computed arithmetic mean.

6) Precision is enhanced by increasing the number of positive samples analyzed.

7) Most systems studied had geometric means of the order of 10^{-1} to 10^{-4} coliforms per 100ml.

8) Most systems studied had geometric standard deviations of the order to 10 to 100 coliforms per 100ml.

9) Systems studied had computed arithmetic means of 0.5 to 700 coliforms per 100ml.

The consequences of these findings on evaluation of bacteriological water quality and the MCL's will be discussed in Section 5.5.

5.4.5 Negative Binomial Distribution

The negative binomial distribution is quite versatile in fitting many dispersion patterns. As stated in Section 5.3.6, the negative binomial has been used to represent coliform dispersion patterns previously (Thomas 1955, Pipes et al. 1977, Muenz 1978, El-Shaarawi et al. 1981). This distribution requires three parameters: mean, variance and k. The parameter k is a coefficient of aggregation and is estimated from the mean and variance. We estimated k using the maximum likelyhood method.

The results of fitting coliform data from the water distribution systems to the negative binomial are shown in Table 5.9. The k's ranged from 0.006 to 0.16. As k decreases the degree of aggregation increases. The fact that all k values were <1 is indicative of a clumped distribution. The analysis for goodness of fit was performed by the Kolmogorov-Smirnov test in which the test statistic is Dmax. Dmax represents the maximum deviation between expected and observed cumulative frequency distributions divided by the sample size. The larger the Dmax the less the probability of expected results fitting the negative binomial. All data fit the negative binomial by this test.

TABLE 5.9

PARAMETERS OF THE NEGATIVE BINOMIAL DISTRIBUTION OF CONFIRMED TOTAL COLIFORMS IN WATER DISTRIBUTION SYSTEMS

System	Mean CFU/100ml	Variance	k	Dmax[a]
CV	0.11	0.67	0.017	0.005
WH I	0.40	3.62	0.028	0.019
II	> 1.52	> 25.21	0.160	0.056
III	> 1.76	> 72.91	0.036	0.025
BB I	0.51	15.41	0.037	0.047
II	> 9.38	> 439.0	0.103	0.082
TP I	> 2.58	> 59.06	0.101	0.041
II	> 6.16	> 274.9	0.090	0.045
DT	> 0.30	> 9.40	0.012	0.009
BL I				
II	> 0.50	> 23.24	0.032	0.032
MI	> 0.40	> 30.92	0.006	0.016
SR	> 0.93	> 25.72	0.072	0.069
MW	0.41	3.41	0.064	0.037
BG	> 2.15	> 69.73	0.056	0.041

[a] The probability of rejecting the hypothesis that the coliforms have a negative binomial distribution in all systems was $p>0.2$.

Thus the dispersion of coliforms in all systems examined could be described by the negative binomial distribution. El-Shaarawi et al. (1981) have discussed the advantages of applying the negative binomial to bacteriological data. However, the estimation of k is laborious and the data are equally well described by the lognormal distribution. As the lognormal distribution is used in other areas of water quality evaluation, we have reasoned that the latter may be a more suitable model upon which to evaluate the MCL's and sampling frequency. Further discussions are restricted to the lognormal distribution.

5.5 RELATIONSHIP OF THE LOGNORMAL DISTRIBUTION TO THE NATIONAL INTERIM PRIMARY DRINKING WATER REGULATIONS

5.5.1 General

There are several consequences of the fact that coliforms are lognormally distributed in water systems to the National Interim Primary Drinking Water Regulations. In this section we discuss these consequences with respect to the following as functions of the number of samples collected:

1) The probability of exceeding the average rule.

2) The probability of exceeding the maximum rule.

3) The probability of obtaining positive samples.

4) The probability of obtaining samples with high densities of coliforms.

5) The probability of exceeding the MCL if the number of samples is reduced.

6) The interpretation of field results when all samples are negative.

The lognormal distribution has several characteristics that should be explained prior to these discussions. Most microbiologists are accustomed to analyses of samples in which the sample mean ($\bar{x}$) is a good representation of the true mean (μ). Such would be the case if coliforms were randomly dispersed. In the lognormal distribution this may not be the case. The sample mean ($\bar{x}$) is associated with the geometric mean (GM), the geometric standard deviation (GSD), and the number of samples (n). The GM and GSD may be used to compute an arithmetic mean [α) ($\ln \alpha = \ln GM + 1/2 (\ln GSD)^2$] which, as has been shown in section 5.4.4, may deviate greatly from $\bar{x}$. Also, $\bar{x}$ is invariably greater than the GM. If coliforms were randomly dispersed, and if a $\bar{x} > 1$ were unacceptable; then a $\mu > 1$ is unacceptable. If coliforms were lognormally dispersed, and if an $\bar{x} > 1$ were unacceptable; then there are several options as to what "true" parameter is used for acceptability. One may use μ, GM, GSD, or α. Another way of stating this is that with a lognormal distribution, various combinations of parameters may be obtained for any $\bar{x}$.

In the following discussions we restrict our examples of parameter combinations. We consider primarily several initial distributions. Two distributions had an initial α equal to 1. The GSD's were 10 and 100 which were values bracketing most of the observed GSD's (Table 5.7). The respective initial GM's were 0.07 and $2.48x10^{-5}$ coliforms per 100ml. From these two distributions calculations for various numbers of samples were made holding either α constant at 1 or holding the GM's constant at either 0.07 or 2.48

x 10^{-5}, respectively. The remaining distributions had a μ of 1 (GM=2.72) or 0(GM=1) and a GSD of either 10 or 100.

5.5.2 Probability of Exceeding the Average Rule

The α for a lognormal distribution is the computed arithmetic mean. If the GM is held constant, as the number of samples used to compute α increases, α approaches GM. The GM is the median and hence the α approaches the median. This is a consequence of the Central Limit Theorem that as the number of samples increases the distribution approaches normality. The GM for the initial distribution is taken as the mean GM ($\overline{GM}$) for different numbers of samples (n) from the population as a whole and the ln GM is equal to the true mean, μ. The average α($\bar{\bar{x}}$) then is a function of μ and the geometric standard error(ln GSE = $\sqrt{\frac{(\ln GSD)}{n}}$). The GSE decreases as n increases. These relationships were used to compute the probability of exceeding the average rule when the GM equaled either 0.07 or 2.48×10^{-5} coliforms per 100ml. The results of these analyses are shown in Table 5.10. The α's can be seen to approach the $\overline{GM}$'s as n is increased. The probability of exceeding the average rule ($\bar{\bar{x}} > 1$) is low under all conditions and approaches 0 with increased number of samples. It would be highly unlikely that the sample mean would exceed 1 under either of these conditions, and the greatest likelihood would be with very small numbers of samples.

TABLE 5.10

PROBABILITY OF EXCEEDING THE AVERAGE RULE AS A FUNCTION OF NUMBER OF SAMPLES WHEN THE GEOMETRIC MEAN IS HELD CONSTANT

Geometric Standard Deviation	10		100	
Geometric Mean (colif./100ml)	0.07		2.48×10^{-5}	
Number of Samples	$\bar{\bar{x}}$	Prob. of $\bar{\bar{x}} > 1$	$\bar{\bar{x}}$	Prob. of $\bar{\bar{x}} > 1$
1	1	0.125	1	0.011
2	0.264	0.051	5.00×10^{-3}	<0.0001
5	0.120	0.005	2.08×10^{-4}	<0.0001
10	0.091	<0.001	7.19×10^{-5}	<0.0001
10	0.091	<0.001	7.19×10^{-5}	<0.0001
20	0.080	<0.001	4.23×10^{-5}	<0.0001
30	0.076	<0.001	3.55×10^{-5}	<0.0001
40	0.075	<0.001	3.25×10^{-5}	<0.0001
50	0.074	<0.001	3.55×10^{-5}	<0.0001
60	0.073	<0.001	2.97×10^{-5}	<0.0001
70	0.073	<0.001	2.90×10^{-5}	<0.0001
80	0.072	<0.001	2.84×10^{-5}	<0.0001

If the μ equals 1, the GM would be 2.72 coliforms per 100ml (μ = ln GM). The α values for two distributions (GSD = 10 and 100) as a function of number of samples are shown in Table 5.11. The $\bar{\alpha}$'s again approach the $\overline{GM}$ as the number of samples increases. The $\bar{\alpha}$'s with few samples are considerably larger than the GM. As the GSD increases, the $\bar{\alpha}$'s increase, of samples required for α to be within a factor of 2 of the GM increases. The probability of exceeding the average rule in these cases is shown in Figure 5.11. The probability that the average is larger than 1 is greater than 50% for all sample sets. When the GSD = 10 the probability of exceeding the average rule is >0.99 when n is greater than 30.

TABLE 5.11

COMPUTED ARITHMETIC MEANS (α) WHEN μ = 1 AS A FUNCTION OF NUMBER OF SAMPLES

Geometric Standard Deviation	10	100
Geometric Mean (colif./100ml)	2.72	2.72
Number of Samples	$\bar{\alpha}$	$\bar{\alpha}$
1	38.51	1.1×10^5
2	10.23	545.5
5	4.62	22.65
10	3.54	7.85
20	3.10	4.62
20	3.10	4.62
30	2.97	3.87
40	2.90	3.54
50	2.87	3.36
60	2.84	3.24
70	2.82	3.16
80	2.81	3.10

The relationship of number of samples to probability of exceeding the average rule when μ or $\overline{GM}$ is held constant is highly dependent on the μ and $\overline{GM}$ chosen. In Table 5.10 the GM's were less than 1, and the probability of exceeding the rule decreased with increased number of samples. In Figure 5.11 the $\overline{GM}$ was greater than 1, and the probability of exceeding the rule increased with increased number of samples. At a GM equal to 1 (hence μ = 0) the probability is always 0.50 and is independent of the number of samples. The α values for two such distributions (μ = 0, GSD = 10 and 100) are shown in Table 5.12. Again, $\bar{\alpha}$ approaches GM as the number of samples is increased, and the $\bar{\alpha}$'s for GSD = 10 decreased more rapidly than those for GSD=100.

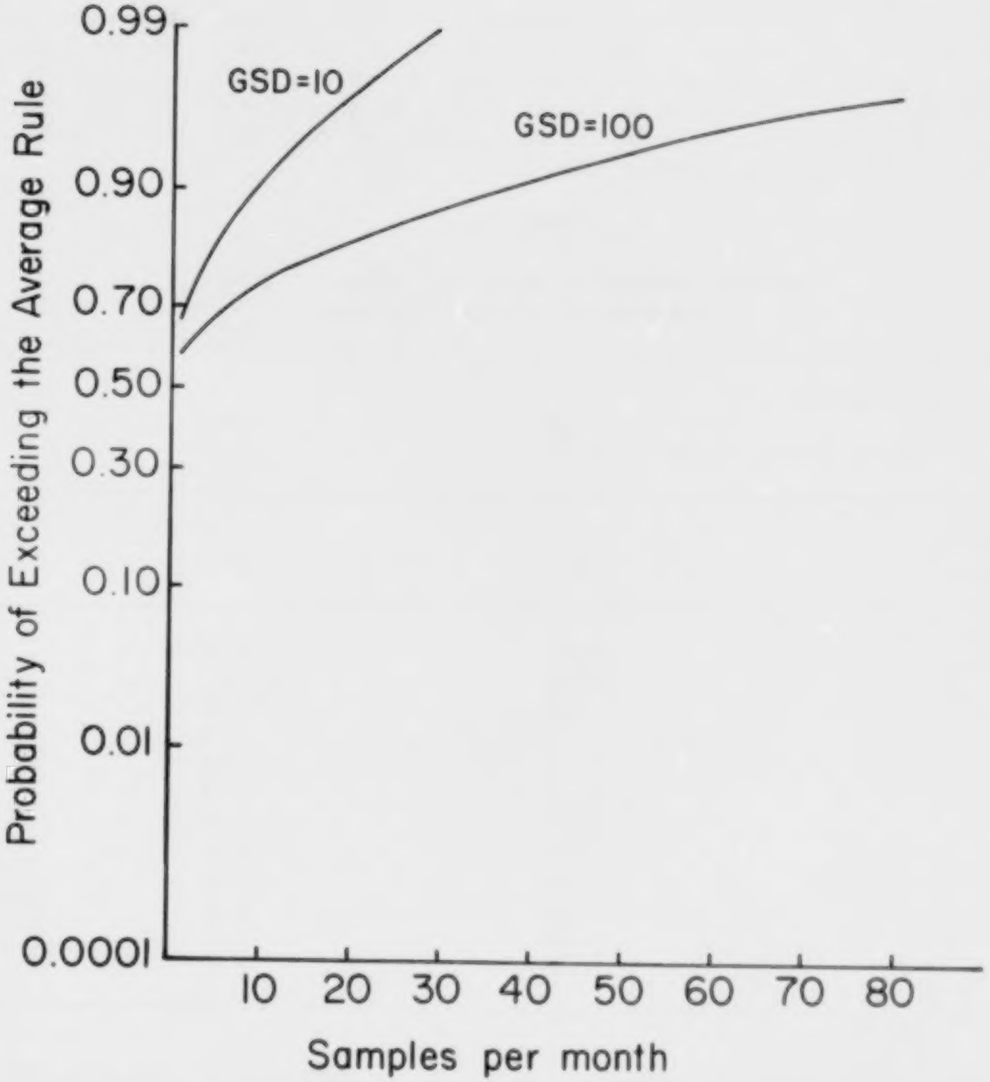


Figure 5.11 Probability of Exceeding the Average Rule when $\mu = 1$ as a Function of Number of Samples

TABLE 5.12

THE AVERAGE COMPUTED ARITHMETIC MEAN (α)
WHEN $\mu = 0$ AS A FUNCTION OF NUMBER OF SAMPLES

Geometric Standard Deviation	10	100
Geometric Mean (colif./100ml)	1	1
Number of Samples	$\bar{\alpha}$	$\bar{\alpha}$
1	14.17	4.03×10^4
2	3.76	200.7
5	1.70	8.34
10	1.30	2.89
20	1.14	1.70
30	1.09	1.42
40	1.07	1.30
50	1.05	1.24
60	1.05	1.19
70	1.04	1.16
80	1.03	1.14

In all of the examples above, μ was held constant. If $\bar{\alpha}$ is held constant, the GM increases to $\bar{\alpha}$ with increased number of samples. The probabilities of exceeding the average rule when $\bar{\alpha} = 1$ are shown in Figure 5.12. The probability of exceeding the average rule increases asymptotically with increased number of samples. The probabilities remained below 0.5 for $n < 80$, and were always higher for GSD = 10 than for GSD = 100.

It should be evident from all of these calculations that the choice of parameter sets is very important to determining the relationship between samples per month and the probability of exceeding the average rule. To summarize, we have examined 4 general cases:

1) After establishment of $\alpha = 1$ and GSD = 10 and 100, the GM's were held constant at 0.07 and 2.48×10^{-5}.

2) The GM was held constant at 2.72 ($\mu = 1$) for GSD = 10 and 100.

3) The GM was held constant at 1 ($\mu = 0$) for GSD = 10 and 100.

4) The $\bar{\alpha}$ was held constant at 1 for GSD = 10 and 100.

In case 1 the probability of exceeding the average rule was very small for all values of n and decreased with increased number of samples. In case 2 the probability approached certainty with increased number of samples. In case 3 the probability was 50% for all numbers of samples. In case 4 the

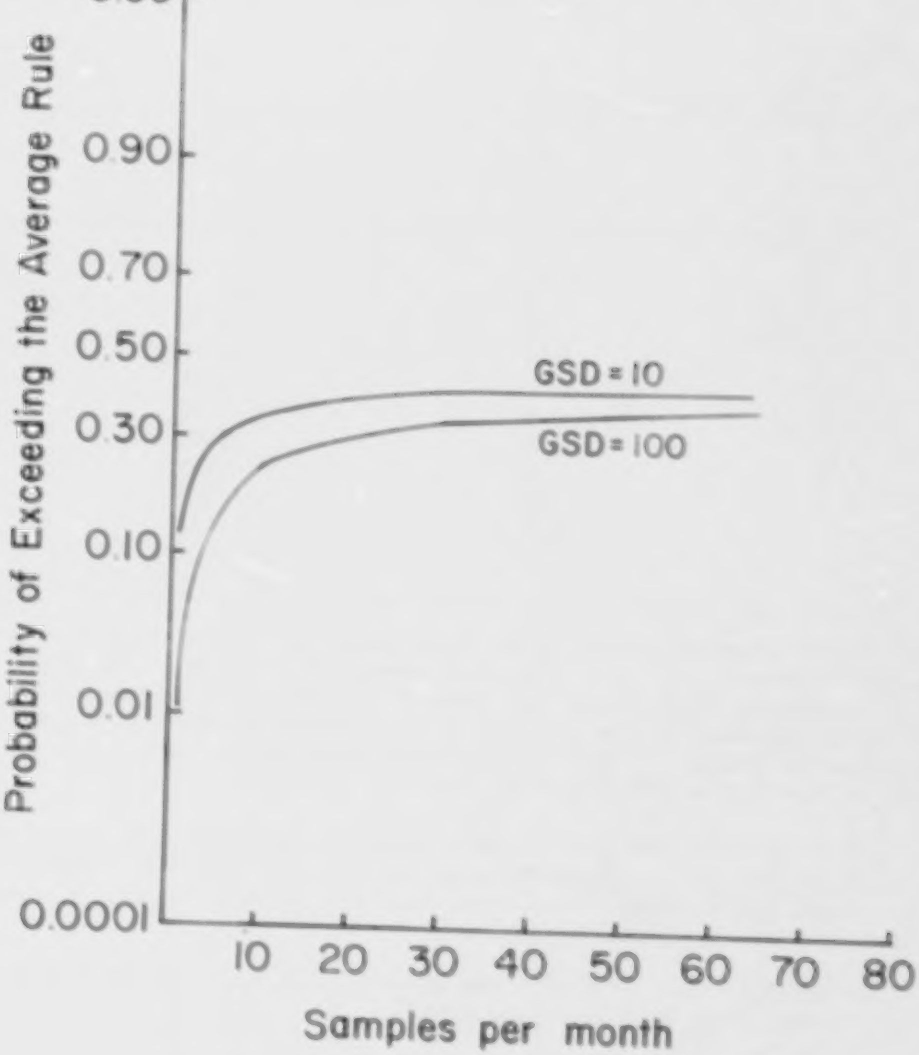


Figure 5.12 Probability of Exceeding the Average Rule when $\alpha = 1$ as a Function of Number of Samples

probability increased with increased numbers of samples approaching 40-42% at 80 samples per month. In all cases where GM was held constant, $\bar{\alpha}$ decreased with increased number of samples, approaching the GM. A sample mean ($\bar{x}$) of >1 may easily be obtained without GM or μ being equal to 1. Thus if it were desirable to establish the basis for regulations on a GM or μ of 1, the amount of bacterial contamination allowable would be greater than current conditions.

In case 1, the GM's were far less than 1 and $\bar{\alpha}$ equaled 1 initially. If these distributions were considered as a basis for the average rule, the probability of $\bar{x} > 1$ would be extremely insensitive to rather large changes in the GM or GSD. For example, if 20 samples were taken per month with a GSD of 10, and the GM rose from 0.07 to 0.54 (a 770% increase) the probability of exceeding the average rule would be only 10%. Thus the average rule is a poor indicator of changes in the geometric mean or median and hence the true mean.

In case 4 the computed arithmetic mean (α) was held constant at 1. The value α embodies both changes in the GM and GSD. For large numbers of samples the GM approaches α. In our field sampling α was greater than 1 coliform per 100ml in all but 3 analyses and greater than $\bar{x}$ in all analyses. Thus exceeding the current average rule would most likely be associated with an $\alpha > 1$. When $n < 10$, the probability of exceeding the average rule increases substantially with increased number of samples. However under the current regulations, if less than 4 samples per month are taken, the results are averaged over a three month period, such that the minimum number of samples for calculation of the average is 3. The probabilities of violation of this minimum would generally range from 10 to 25%. In other words if α equaled 1 and 1 sample per month was averaged over three months the system would be in violation of the average rule once in every 4 to 10 quarters. But if 3 samples were taken per month and averaged per quarter, violation would be very likely to occur at least once per year. When $n = 6$ per month, the rule would be exceeded about twice per year, enough to insure that remedies would be taken to improve bacteriological water quality. For systems taking 10 or more samples per month the probability of violation is at least 0.25 or at least once in every 4 months.

5.5.3 Probability of Exceeding the Maximum Rule

The probability of exceeding the maximum rule is a function of the number of samples and the parameter sets chosen for the lognormal distribution. In this section we consider two cases where: 1) the initial $\alpha = 1$ and the geometric means are held constant at 0.07 for GSD = 10 and 2.48×10^{-5} for GSD=100, and 2) the α is held constant at 1 but the initial conditions are set as above. The probability of exceeding the maximum rule <0.0001 in the first case (Table 5.13). Violation would occur about 1 month every 833 years. Clearly, the rule is insensitive to such distributions when the GM is far less than 1 coliform per 100ml.

The probability of violation for the second case is shown in Figure 5.13. The probabilities of violation for both distributions increase with increased numbers of samples from 2 to 39 samples for which 2 samples with >4 coliforms per 100ml exceeds the rule. Thus as the number of samples increases the probability of obtaining 2 such samples increases. Above 39 samples 1 more sample of >4 per 100ml is required for each 20 samples added. Thus the probability of obtaining 3 such samples in 40 is less than obtaining 2 in 39. The sawtooth effect seen here is similar to that seen for the Poisson distribution (Figure 5.2).

TABLE 5.13

PROBABILITY OF EXCEEDING THE MAXIMUM RULE AS A FUNCTION OF NUMBER OF SAMPLES WHEN GEOMETRIC MEAN IS HELD CONSTANT.

GSD GM (colif./100ml)	10 0.07	100 2.48×10^{-5}
Number of Samples	Probability of Violation	Probability of Violation
2	<0.0001	<0.0001
5	<0.0001	<0.0001
10	<0.0001	<0.0001

The maximum probability of violation occurs at 39 samples for both distributions. The peaks of the sawtooths decrease with increased number of samples. The maximum probability of violation for GSD=10 is 0.53 and for GSD=100 is 0.013. In other words with these distributions violation would occur no more than 1 month in approximately 2 for GSD=10 and 1 month in approximately 6.4 years for GSD=100. As the GSD reflects the degree of aggregation of coliforms, the probability of exceeding the maximum rule decreases as aggregation or patchiness increases. This rule is very insensitive to very patchy distributions of coliforms.

The maximum rule is generally less sensitive to distributions studied than is the average rule. The probability of exceeding the average rule was greater than that of exceeding the maximum rule with few exceptions. They all occurred when α was held constant for a GSD= 10 at n = 36 to 39 and n = 58 to 59. The differences between the two rules at these exceptions were not substantial. Thus, the maximum rule is redundant to the average rule, and in almost all circumstances violation of the maximum rule would be accompanied by violation of the average rule.

This concept was tested with field data shown in Table 5.14. The mean density per location was used for evaluation. Of the 15 analyses 6 exceeded the average rule, and 7 exceeded the maximum rule. In no case the average rule but not the maximum rule was violated, and in 1 case was the maximum rule violated without also violating the average rule. These results are from large numbers of samples. Can the same be said for small numbers of samples? Nine sampling dates from WH are analyzed independently for probability of exceeding either rule (See appendix B). The number of samples

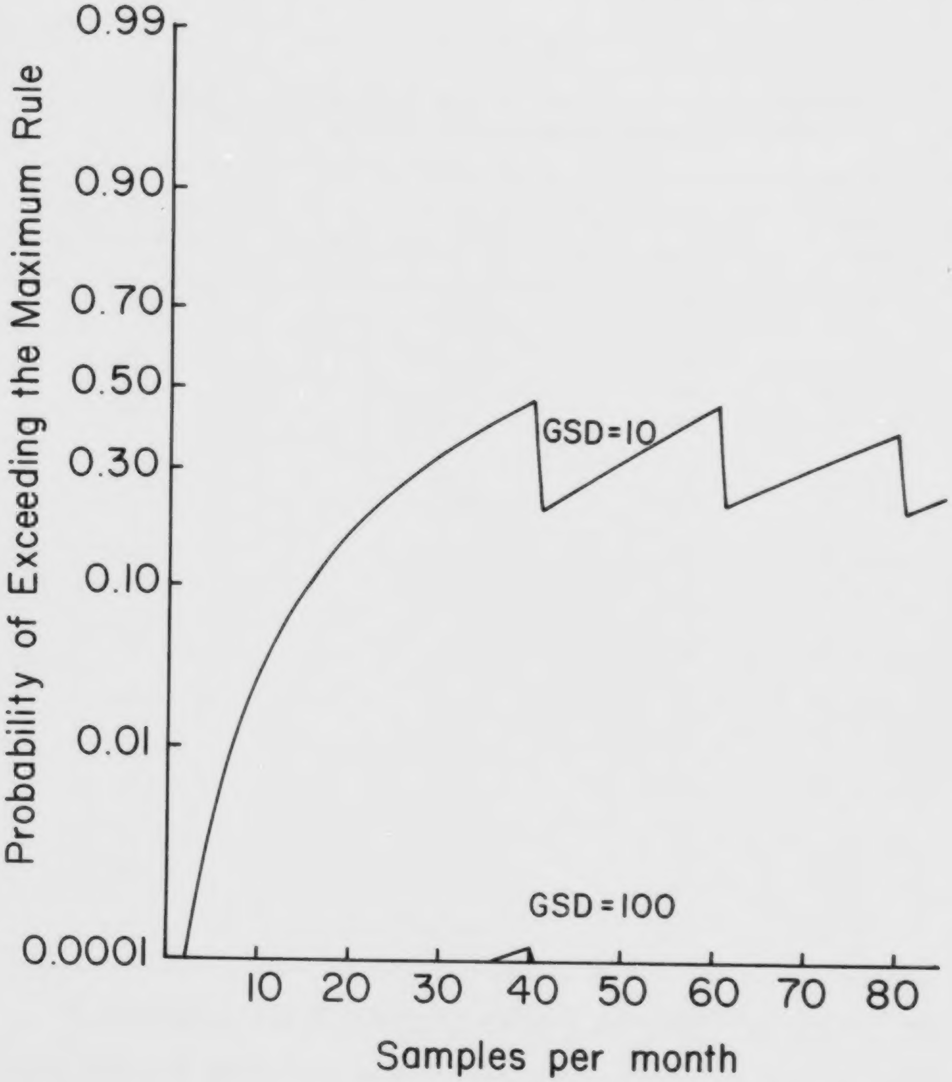


Figure 5.13 Probability of Exceeding the Maximum Rule when α = 1 as a Function of Number of Samples

ranged from 6 to 50 per day. On 5 days the average rule was exceeded, and the maximum rule was violated on 4 of these days. At no time was the maximum rule violated without violation of the average rule. Thus the theoretical relationship between the maximum and average rules is generally confirmed by our field data. It should be noted that these analyses were based on locations from which volumes greater than 100ml were filtered.

TABLE 5.14

OCCURRENCE OF VIOLATIONS IN THE AVERAGE AND MAXIMUM RULES IN WATER DISTRIBUTION SYSTEMS STUDIED

System	Sample Mean	Violation of Average Rule	% >4 colif.per 100ml	Violation of Maximum Rule
CV	0.11	No	1.3	No
WH I	0.40	No	4.5	No
II	> 1.52	Yes	10.4	Yes
III	> 1.76	Yes	5.9	Yes
BB I	0.51	No	2.2	No
II	> 9.38	Yes	23.2	Yes
TP I	> 2.58	Yes	11.5	Yes
II	> 6.16	Yes	18.5	Yes
DT	> 0.30	No	1.1	No
BL I	> 0.35	No	0.4	No
II	> 0.50	No	0.6	No
MI	> 0.40	No	1.4	No
SR	> 0.41	No	4.9	No
	0.41	No	5.1	Yes
BG	> 2.15	Yes	10.9	Yes

5.5.4. Probability of Obtaining Samples above Particular Densities

In this section we discuss the likelihood of obtaining positive samples and samples of particular density ranges in water distribution systems. From our field data and retrospective studies, two points are clear. First, most samples from water distribution systems under normal conditions will be devoid of coliforms. In Table 5.4 we show the fraction of locations for each system studied that produced samples without coliforms. In all systems but two that met the average rule greater than 90% of the samples (locations) had 0 coliforms per 100ml. Even when the average rule was violated over one half of the locations were without coliform growth. Second, when positive locations are obtained the likelihood of obtaining a high density count is relatively large and the likelihood of obtaining a sample that is too numerous to count (TNTC = 80 per 100ml) is finite. TNTC samples were obtained in 11 system analyses. Of these the average rule was violated only 6 times. What then are the consequences of the lognormal distribution on obtaining positive samples and high density samples?

The two cases of theoretical parameter sets discussed in the previous section will be used here. When GM was held constant at 0.07 (GSD=10, $\alpha = 1$) the probability of obtaining at least 1 positive sample was 0.125 for 1 sample, 0.099 for 2 samples, 0.025 for 5 samples and <0.001 for any number above 10 samples. When the GM was held constant at 2.48×10^{-5} (GSD 100, $\alpha = 1$), the probability was 0.010 for 1 sample and <0.0002 for all higher samples. Thus the probability of obtaining a positive sample is very small for low GM's when GSD ranges from 10 to 100.

When α is held constant at 1 and GSD's are 10 and 100, the probability of obtaining positive samples increases with increased number of samples (Figure 5.14). The probability of obtaining a positive sample increases to greater than 99% for $n > 40$ when GSD=10. The probability is always less when GSD=100 than for GSD=10 and reaches 50% by 63 samples. If GSD=10, at least one positive sample would be found on average <2 months if 6 or more samples are taken per month. If GSD=100, at least 63 samples would be required for the same result. Thus as the coliforms become more clumped in their distribution, the likelihood of obtaining positive samples decreases.

If the likelihood of obtaining positive samples is very small when GM's are held constant and are small, the probability of samples with high coliform densities is even smaller. The discussion of samples with high coliform densities is restricted to the cases where α is held constant at 1. The probabilities of obtaining at least 1 sample with 4 or more coliforms per 100ml are shown in Figure 5.15. Again probabilities increase with increased number of samples and are always greater for GSD=10 than GSD=100. These probabilities are similar to but not exactly those of check sampling. The density of samples for which check sampling is required is actually >4 coliforms per 100ml. However, this depiction is a reasonable guide. If GSD=10 and $\alpha = 1$, a check sampling program would occur at least once every two months when 15 or more samples are taken and at least once every four months when >6 samples are taken per month. If GSD=100 and $\alpha = 1$, check sampling would occur for less often (for example once every 14 months when 15 samples are taken per month).

The occurrence of a TNTC sample is often taken as an error in sampling or processing by water system personnel. A TNTC count is particularly damaging to the compliance to the average rule for small water systems. It is often considered as an 80 coliforms per 100ml density; and thus when one TNTC is found, 79 negative samples must be obtained if the average rule is to be met. What then is the probability of obtaining a TNTC? Is the probability high enough such that obtaining a TNTC may not be regarded solely as an error? The probabilities for obtaining at least 1 TNTC, when $\alpha = 1$ and GSD's=10 and 100, are shown in Figure 5.16. The probabilities for neither distribution exceed 10% for $n < 80$ per month. However, when the number of

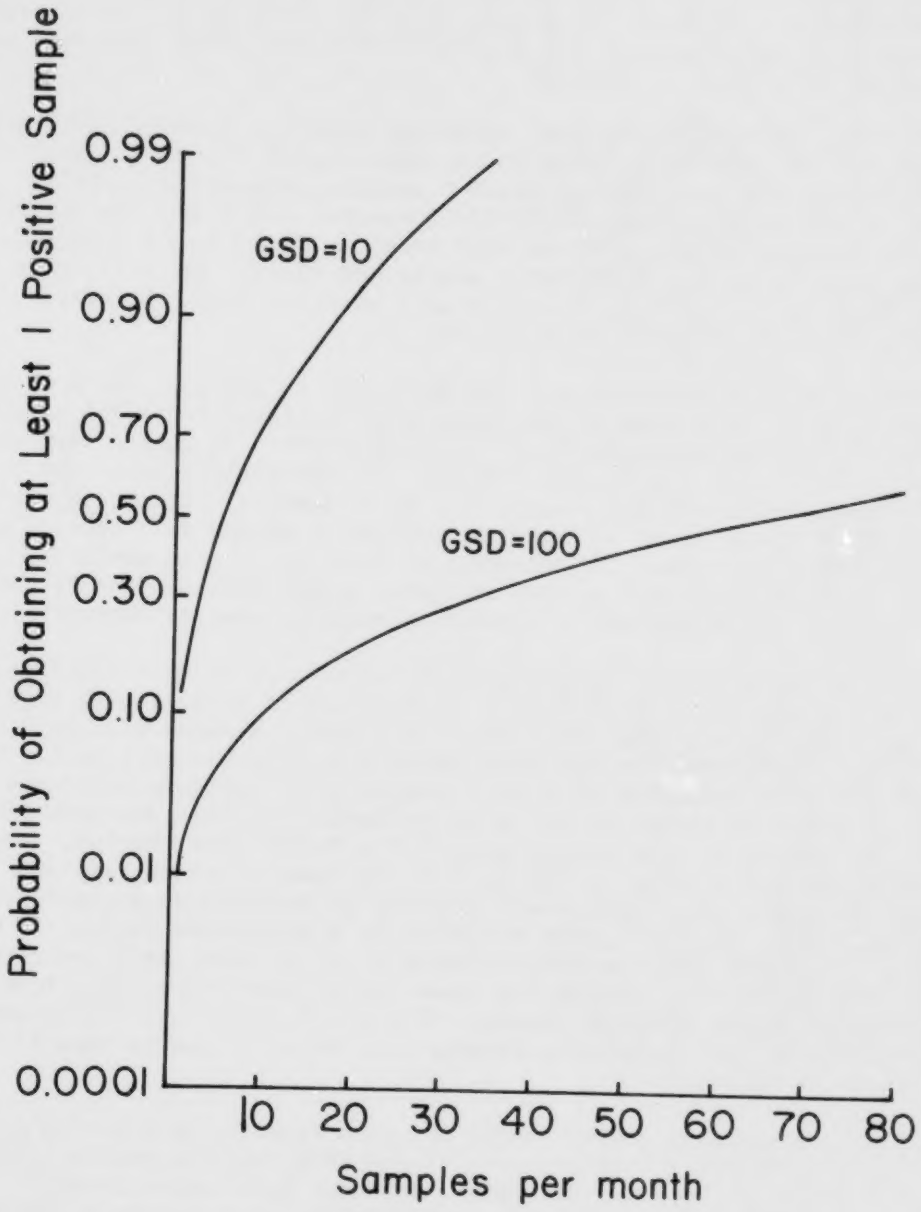


Figure 5.14 Probability of Obtaining at Least 1 Positive Sample as a Function of Number of Samples when $\alpha = 1$.

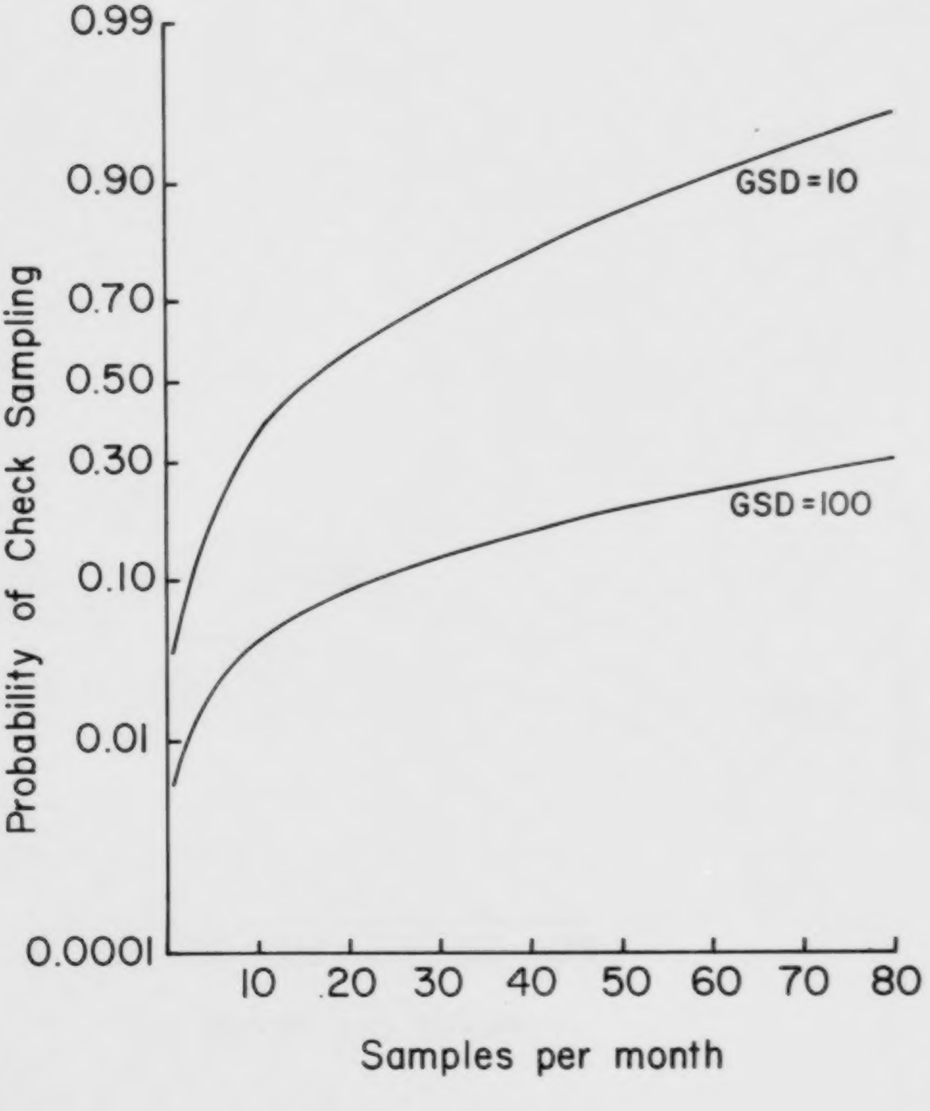


Figure 5.15 Probability of Obtaining at Least 1 Sample with ≥ 4 Coliforms per 100ml when α = 1 as a Function of Number of Samples

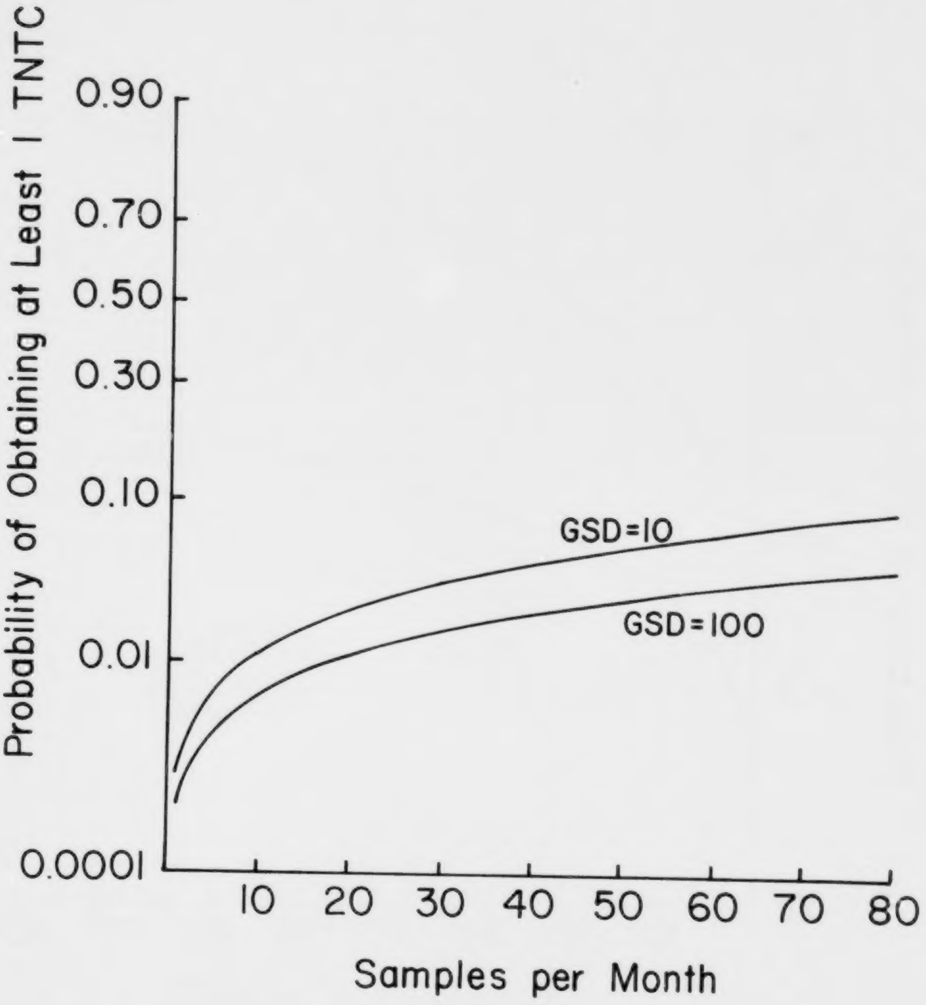


Figure 5.16 Probability of Obtaining at Least One TNTC Sample when α = 1 as a Function of Number of Samples

samples is greater than 9 for GSD=10 and 19 when GSD=100 the probability is greater than 1%. For most systems having an $\alpha = 1$ and sampling 20 times per month, a TNTC may be expected to be found at least every 8 to 10 years. With this low probability it can be seen why most operators consider such samples as errors. However, another way of looking at this aspect is to consider that if 100 systems sample 20 times per month with $\alpha = 1$ for all, at least 1 system would probably have a TNTC. The state or federal official examining the records of many systems on a monthly basis must be prepared for the eventuality of occasional TNTC samples.

5.5.5. The Probability of Exceeding Regulations if the Number of Samples is Reduced

The National Interim Primary Drinking Water Regulations (U.S.E.P.A. 1976) state that as many as 75% of the microbiological samples required per month may be substituted by free chlorine measurements with 4 chlorine measurements required per coliform determination (Section 141.21h). As the likelihood of violating the average and maximum rules are dependent on sample size, one must consider the consequences of such a rule. Our discussion is restricted to the cases where α is held constant at 1 coliform per 100ml and GSD=10 and 100.

The probability of exceeding the average rule would largely be unaffected if the number of microbiological samples examined is >20 per month. When sampling falls below 20, then the probability of exceeding the average rule is diminished significantly as fewer samples are taken. Clearly, this rule should be given special scrutiny for small water systems.

The sawtooth nature of the probability of exceeding the maximum rule should be recognized in establishing the degree of substitution. The probability of exceeding the maximum rule can change by more than 0.2 by a difference of 1 sample (e.g., 39 to 40 samples in Figure 5.13). Substitutions may be made such that the remaining number of samples examined for coliforms maximizes the sensitivity of sampling within each 20 sample unit. That is the retained microbiological samples could be set at 39, 59, 79, etc. If the retained microbiological sample size falls below 39, the probability of exceeding the maximum rule decreases significantly for fewere samples. Thus the cholorine substitution rule can be used by small systems to avoid violation of the MCL even when microbiological contamination is high.

5.5.6. Interpretation of Field Results When All Samples are Negative

What can be said about the degree of bacterial contamination of a water system if all samples in a month are negative? From the previous discussions it can be seen that systems may be contaminated with that contamination rarely being detected. This was certainly the case when the GM's were held constant at 0.07 and 2.48×10^{-5} coliforms per 100ml and GSD's were 10 and 100 respectively. Also, in spite of the fact that we found coliforms in every system we studied, the historical records of those systems rarely noted coliforms. In fact in system LB no coliforms were reported for several years prior to our study and were not obtained by required monitoring during our study. The possibility existed that the procedures used by those at LB were not satisfactory for proper recovery of coliforms (see Section 4). But the fact remains that coliforms may occur in relatively high densities in a system and go undetected.

The probability of α being 1 and all samples during a month being negative (0 coliforms per 100ml) may be derived from Figure 5.14. As one minus the probability of obtaining at least one positive sample. For example if 20 samples are taken in one month and all are negative, the probability of this occurring when $\alpha = 1$ and GSD=100 is 0.8. In other words 80% of the months that 20 samples are taken, detection of coliforms would not occur even though they would be present with the above distribution. When $\alpha = 1$ and GSD=10, contamination should be detectible in the form of at least 1 positive sample at least half the time when greater than 5 samples are taken per month. If only one sample is taken per month 7 out of 8 months would be negative. The probability of detecting coliforms decreases as the GSD increases. If only one sample is taken per month when $\alpha = 1$ and GSD=100, 90 out of 91 months would have no positive samples. Even when 40 samples are taken per month approximately two out of three months may not give positive samples. Thus, small numbers of samples, when coliform distributions are highly aggregated, are quite insensitive to the detection of contamination.

5.6 REPLICATION OF SAMPLES AND LARGER VOLUME SAMPLES

5.6.1 Introduction

During the sampling programs it was found that the time required per site sampled varied from about 10 minutes to 30 minutes because of the probability of finding someone at home to provide access to a tap, and the number of refusals to permit access to the residence. The amount of time needed to measure water temperature and chlorine residual and obtain the sample was small in comparison with the time required to gain access to the tap. Thus, the number of samples collected on any given day could be readily increased with little additional time and effort by obtaining replicate samples at each site sampled.

Replicate samples at a site are not independent observations on the quality of water in the system and thus replication of samples is not equivalent to obtaining additional samples at other sites. However, it is clear that obtaining replicate samples at each site provides more information than single samples at each site. How much more information is obtained from the second and additional replicates is the subject of this Section.

One method of interpreting the data from replicate samples is to consider them as equivalent single, larger volume samples. Two replicate 100ml samples could be considered to be the same as a single 200ml sample or two replicate 200ml samples as a single 400ml sample and so forth. Larger volume samples, mostly 1000ml and 200ml samples, were collected from some communities on some sampling days to test this interpretation.

The purpose of this Section is to explore the questions associated with replication of samples from individual sites and with the use of larger volume samples. The questions are concerned with the most efficient method for monitoring a water distribution system. The objective of bacteriological monitoring is to find coliform organisms if they are present or, stated from the opposite viewpoint, to demonstrate that the disinfection of the water during treatment and the integrity of the distribution system are maintained so that the probability of bacterial contamination of the water is very small. Replication of samples and/or use of larger volume samples are evaluated in terms of how they might expedite the achievement of these objectives.

5.6.2 Replicate Samples Collected

The number of sites sampled and replicate samples collected for the various communities included in this project are given in Tables 5.15-5.23 along with the confirmed coliform counts in samples where they were present. The numbers assigned to the individual replicates represents in most instances, the order of collection.

One liter samples were collected during the sampling of systems CV and WH. It was found that if large amounts of particulate matter were present in the samples, filtration of 1000ml required an excessive amount of time (30 to 60 minutes), and the particulate matter on the surface of the filter interfered with development of typical coliform colonies. The use of one liter samples was discontinued about half way through the WH system sampling.

During the second half of the CV system sampling and thereafter a minimum of two 100ml replicates were collected at each site sampled. In DT and BL some extra 100ml samples were collected at a few sites and a few 200ml samples were also collected. In MI, SR, MW and BG duplicate 200ml samples were frequently collected in addition to the two 100ml replicates. One liter replicates were tried again on two sampling days in MI because the water there was exceptionally free of particulate matter, but no coliforms were found in any of those samples.

5.6.3 Results from Duplicate Samples

The largest mass of data on replicate samples which we have collected consists of counts on duplicate 100ml samples. These data are summarized in Table 5.24. All sites where less than 2 replicate 100ml samples were collected are not included. For sites where more than 2 samples were collected data from only the first two 100ml samples are used.

One or more samples with more than eighty coliforms per 100ml were collected from each of the systems sampled except CT, WH, and MW. When one or more coliform results are recorded as TNTC the calculation of a mean coliform density is uncertain. To obtain the figures in Table 5.24, TNTC results were taken as more than eighty coliforms per 100ml. This gives a minimum value for the mean and variance of the density.

From the confidence intervals in Table 5.24 it can be readily seen that there are no significant differences in the fraction of samples with coliforms or in the mean density between the first and second replicates. This is as expected and clearly shows that the chance of finding a coliform is

TABLE 5.15

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
COATESVILLE

Sampling Date	Sites Sampled	Nos. of 100ml Replicates	Nos. of 1000ml Replicates	Sites without Coliforms	Sequences of Coliform Counts at Positive Sites
2/1/79	16	1, 3	2[a]	16	
2/6/79	24	1	2	22	1-0, 9-0
2/9/79	16	1	2[b]	15	3-0
2/14/79	10	2, 3, 4	1	10	
2/15/79	21	1, 2, 3	4	20	32-0-0-0
3/6/79	34	2, 3	1	31	T-0-0, 1-0-0, 1-0-0
3/8/79	30	2, 3	1	26	13-0-0, 0-2-0, 1-0-0, 7-0-0
3/13/79	33	2, 3	1	30	0-0-1, 0-2-0, 1-0-4
3/15/79	41	2, 3	1	39	?[c]-1-0, ?[c]-9-0

---

a. Volume filtered was 800ml rather than 1000ml.

b. One liter filtered in separate portions of 100ml, 200ml,and 700ml.

c. After incubation for 22 hours on M-Endo medium, the filter from the 1 liter sample was placed in lauryl tryptose broth and produced gas in 24 hours of incubation.

TABLE 5.16

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
WOODBURY HEIGHTS

Sampling Date	Sites Sampled	Nos. of 100ml Replicates	Nos. of 1000ml Replicates	Sites without Coliforms	Sequences of Coliform Counts at Positive Sites
4/24/79	6	2, 3	1	6	
4/26/79	17	2, 3	1	14	6-0-13, ?[a]-11-16, ?[a]-0-1
5/1/79	22	2, 3	1	20	2-0-0, 1-0-0
5/3/79	21	2, 3	1	18	0-0-1, T[b]-0-0, 1-0-9
5/8/79	13	1, 2	-	6	1-0, 1-0, 1-26, 11-1, 3-0, 0-2, 6-0
5/10/79	50	1, 2	-	35	0-11, 3-1, 1-0, 2-1, 0-1, 0-1, 9-47, 3-3, 0-3, 0-3, 2-1, 0-1, 1-1, 1-3, 0-3
5/15/79	30	1, 2	-	19	2-0, 0-2, 12-0, 1-1, 0-1, 1-35, 0-2, 3-0, 4-0, 1-0, 2-7
5/16/79	6	1, 2	-	6	
5.17/79	50	1, 2	-	25	11-0, 1-0, 1-0, 1-0, 1-0, 0-3, 0-3, 5-3, 1-2, 0-18, 0-1, 1-0, 48-46, 0-15, 0-1, 1-0, 3-24, 22-0, 0-1, 1-0, 9-0, 1-0, 1-0, 1-0, 0-8

---

a. After incubation for 22 hours on M-Endo medium, the filter from the 1 liter sample was placed in lauryl tryptose broth and produced gas in 24 hours of incubation.

b. T = too numerous to count = more than 80 colonies per filter.

TABLE 5.17

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
LONG BEACH WATER CO.

	BRANT BEACH PLANT SERVICE AREA			TERRACE PLANT SERVICE AREA		
Sampling Date	Sites Sampled	Sites Without Coliforms	Sequences of Counts of Positive Sites	Sites Sampled	Sites Without	Sequences of Counts of Positive Sites
6/7/79	1	1		2	2	
6/12/79	17	17		24	24	
6/14/79	24	17	0-1,23-52,0-2,0-1, 0-1,1-0,0-1	33	17	7-14,0-60,1-4,0-1,0-T, 0-43,0-1,0-1,0-1,0-54, 2-0,0-T,0-1,0-1,0-T,0-3
6/19/79	21	19	2-0,1-0	29	11	2-0,2-0,2-3,4-8,2-0, 2-2,0-1,8-0,0-2,1-0, 0-1,0-18,1-2,3-2,3-16, 25-9,5-29,21-0
6/21/79	29	27	1-0,7-1	34	27	0-5,5-1,1-1,0-4,0-6, 0-1,1-0
7/10/79	25	18	0-2,10-4,26-0,0-1, 1-0,T-4,50-T	31	20	0-5,3-1,T-77,5-T,0-T, 0-T,1-0,8-3,0-6,1-71, 45-0
7/11/79	6	6		6	6	
7/12/79	20	6	T-T,0-T,0-14,T-T,T-T, T-8,T-1,14-0,5-0,0-T, 0-2,0-T,54-T,1-T	36	15	8-0,3-6,1-0,1-T,2-0, 0-1,0-T,0-T,1-1,T-3, 1-0,T-T,2-0,0-50,T-T, T-T,14-T,T-0,7-0,5-4, 0-33,0-1
7/17/79	14	5	1-0,0-1,0-10,3-6,0-4, T-T,0-1,26-T,0-1	22	14	10-1,2-0,15-0,2-0,3-0, 1-0,0-2,6-0
7/18/79	14	8	2-0,0-3,9-T,10-0,1-1, 3-1	22	15	49-60,2-4,3-1,27-0,26-0, 2-0,0-1
7/19/79	14	9	4-0,0-3,3-39,6-0,3-0	23	17	1-0,3-4,1-0,0-2,0-1,1-0

T= too numerous to count = more than 80 colonies per filter.

TABLE 5.18

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
DOWNINGTOWN

Sampling Date	Sites Sampled	Nos. of 100ml Replicates	Nos. of 200ml Replicates	Sites Without Coliforms	Sequences of Coliform Counts at Positive Sites
12/27/79	5	1,2	-	5	
1/8/80	9	1,2	-	8	0-1
1/15/80	28	1,2	3,4	24	0-0-1-0,0-0-2-0,0-0-1-0,1-0-0-0
1/16/80	4	1,2,3,4	-	4	
1/17/80	3	1,2,3,4	-	3	
1/22/80	22 2	1,2 1	-	22 2	
1/23/80	4	1,2	-	4	
1/24/80	3	1,2	-	3	
1/28/80	8	1,2	-	8	
1/29/80	19 5	1,2 1,2,3,4	- -	15 5	1-0,4-0,0-11,T[a]-0
2/5/80	17 5	1,2 3,4	- 1,2	16 5	1-1
2/12/80	42	1,2	-	41	0-5

a. T = too numerous to count = more than 80 colonies per filter.

TABLE 5.19

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
BROOKLAWN

Sampling Dates	Sites Sampled	Nos. of 50ml Replicates	Nos. of 100ml Replicates	Nos. of 200ml Replicates	Sites Without Coliforms	Sequences of Coliform Counts at Positive Sites
3/4/80	12	-	1, 2	-	11	0-1
3/11/80	29	-	1, 2	-	29	
3/12/80	6	-	1, 2	-	6	
3/17/80	6	-	1, 2	-	6	
3/18/80	24	-	1, 2	-	23	1-0
3/19/80	6	-	1, 2	-	6	
3/21/80	7	-	1, 2	-	6	0-1
3/25/80	23	-	1, 2	-	23	
3/26/80	8	-	1, 2	-	8	
3/27/80	8	-	1, 2	-	8	
4/1/80	6	-	1, 2	-	6	
4/2/80	8	-	1, 2	-	7	5-7
4/7/80	8	-	1, 2	-	8	
4/8/80	30	-	1, 2	-	30	
4/15/80	20	-	1, 2	-	17	T[a]-T, 0-1, 1-0
4/16/80	8	-	1, 2	-	8	
4/22/80	22	-	1, 2	-	22	
4/23/80	2	-	1, 2, 3, 4, 5, 6	7, 8	1	0-1-1-0-0-0-0-0
4/24/80	6	1, 2, 3, 4	5, 6, 7, 8	-	6	

a. T = too numerous to count = more than 80 colonies per filter.

TABLE 5.20

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
Mt. Idy Mobile Home Park

Sampling Date	Sites Sampled	Nos. of 100ml Replicates[a]	Nos. of 200ml Replicates[a]	Nos. of 1000ml Replicates	Sites Without Coliforms	Sequences of Coliform Counts at Positive Sites
5/6/80	25	1,3	2,4	-	24	0-0-1-0
5/8/80	24	2,3	1,4	-	23	1-0-0-1
5/13/80	19	1,3	2,4	-	14	0-3-0-0, 0-1-0-0, T-T-T-T[b], 0-1-0-0, 1-0-1-3
5/15/80	20	2,4	1,3	-	20	
5/20/80	24	1,4	2,3	-	24	
5/22/80	25	1,3	2,4	-	25	
5/27/80	23	3,4	1,2	-	23	
5/29/80	21	1,2	3,4	-	20	0-0-0-1
6/3/80	14	2,5	1,4	3	14	
6/5/80	9	3,4	1,2	3	9	

a. Order of collection of 100ml and 200ml replicates varied on individual sampling days but the order cited representes the volumes of samples with coliforms.

b. T = too numerous to count = more than 80 colonies per filter.

TABLE 5.21

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
SPRING RUN

Sampling Date	Sites Sampled	Nos. of 100ml Replicates	Nos. of 200ml Replicates	Sites with no coliforms	Sequences of Coliform Counts at Positive Sites
9/25	6	-	1,2⁻	6	
10/8	37	1,2	-	35	0-1, 0-1
10/16	31	1,2	-	31	
10/21	25	1,2	3,4	16	0-2-0-0, 0-0-1-0, 0-1-0-0, 0-0-1-0, 0-0-1-0, 0-0-1-0, 3-6-0-0, 12-0-1-0, 0-0-0-1
10/23	18	1,2	3,4	4	0-0-1-0, T-0-0-0, 0-11-1-1, 0-0-0-1, 0-0-1-14, 1-0-0-0, 0-0-0-1, 1-0-0-0, 0-0-0-2, 0-0-8-0, 6-15-0-8, 0-0-0-9, 45-9-6-2, 0-7-12-1
10/28	12	1,2	3,4	8	0-0-0-1, 7-6-T[a]-T, 1-3-0-0, 0-T-45-38
10/30	9	1,2	3,4	8	0-0-1-0
11/5	21	1,2	3,4	11	1-0-0-0, 1-0-0-0, 6-0-31-2, 0-2-0-0, 0-0-0-1, 0-2-0-0, 0-0-1-0, 1-0-0-0, 0-0-0-?[b], 0-0-?-?
11/6	5	1,2	3,4	4	10-1-0-0
11/11	8	1,2	3,4	8	
11/13	4	1,2	3,4	3	0-0-1-0
11/19	4	3,4,7,8	1,2,5,6	4	
12/2	6	3,4	1,2	5	0-7-0-0

a. T = too numerous to count = more than 80 colonies per filter

b. After incubation of 22 hours on M-Endo medium, the filters from some 200ml samples were placed in lauryl tryptose broth and produced gas in 24 hours of incubation.

TABLE 5.22

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
MARSHALLTON WOODS

Sampling Date	Sites Sampled	Nos. of 100ml Replicates	Nos. of 200ml Replicates	Sites Without Coliforms	Sequences of Coliform Counts at Positive Sites
10/8/80	4	1,2	-	4	
10/23/80	4	1,2	3,4	2	1-0-1-1, 0-1-0-0
10/30/80	13	1,2	3,4	9	0-1-0-0, 2-0-0-0, 2-0-0-0, 0-0-0-1
11/6/80	19	1,2	3,4	16	0-0-19-40, 8-0-0-0, 0-0-6-0
11/11/80	15	1,2	3,4	13	1-0-0-0, 0-0-26-0
11/13/80	13[a]	1,2,5,6	3,4,7,8	6	63-0-0-2-5-0-1-1, 0-0-0-0-0-3-3-0, 0-0-0-0-0-0-0-2, 14-0-0-0-0-0-0-0, 1-0-0-2-0-1-0-7, 1-0-0-0-0-0-0-0, 1-12-0-0-0-0-0-0
	1	3,4	1,2,5,6	0	1-0-1-1-0-0
11/19/80	14	3,4,7,8	1,2,3,4	14	
12/11/80	16	3,4,7,8	1,2,5,6	14	0-0-0-3-0-0-0-0, 0-0-0-1-0-0-0-0

a. More than eight replicate samples were collected at some sites and the order of collection of 100ml and 200ml replicates was varied. The order cited represents the volumes of samples with coliforms.

TABLE 5.23

SEQUENCES OF COLIFORM COUNTS IN REPLICATE SAMPLES
BRADFORD GLEN

Sampling Date	Sites Sampled	Nos. of 100ml Replicates	Nos. of 200ml Replicates	Sites With No Coliforms	Sequences of Coliform Counts at Positive Sites
9/25/80	5	-	1,2	3	4-0, 0-1
10/8/80	4	1,2	-	4	
10/16/80	13	1,2	-	13	
10/21/80	3	1,2	3,4	2	T[a]-0-0-0
10/28/80	3	1,2	3,4	0	2-0-59-6, 7-1-T-0, 0-T-8-T
10/30/80	4	1,2	3,4	2	8-1-0-0, 2-0-0-8
11/20/80	7	3,4,7,8	1,2,5,6	5	T-4-0-0-0-0-0-0, 13-9-0-2-1-0-0-0
12/2/80	7	1,2	3,4	7	

T = too numerous to count - more than 80 colonies per filter the same for either replicate. The only physical difference between the two replicates is the length of time the water was allowed to run from the tap before the sample was collected, and this is not expected to influence the chance of finding a coliform.

The important question about the use of replicate 100ml samples is concerned with how much is gained by the collection of replicate samples at one site as compared with the collection of single samples at other sites. Empirical attempts to develop an answer to this question were made during the sampling of the systems LB, DT, BL, SR, MW and BG.

5.6.4. Negative Sampling Results

On June 14, 1979 twenty-nine replicate 100ml samples were collected at a single residence in the BB service area. Two of the samples had presumptive coliforms but they did not verify. In the rest of the BB service area on that day, 7 out of 24 sites sampled had coliforms present in one or both of the samples collected, and all colonies picked verified as coliforms. On July 19, 1979 nineteen replicate samples were collected in a motel room in the TP service area but none of them had coliforms present. In the rest of the TP service area on that day 6 out of 22 sites sampled had verified coliform organisms present. The conclusion from this is that even when coliform densities are relatively high in a good portion of a distribution system there can be parts of the system where coliforms are absent or, at most, present in very, very low densities.

During the sampling of DT 4 replicates were collected at 4 sites on 1/16/80, at 3 sites on 1/17/80, at 5 sites on 1/29/80 and at 5 sites on 2/5/80. None of these samples had coliforms present. During the sampling of BL 8 replicate samples were collected at 6 sites on 4/24/80 but no coliforms were found. In MI 4 replicates were collected from each site sampled and no coliforms were found on six different sampling days on which a total of 115 sites were sampled. In SR no coliforms were found in 4 replicates from 8 sites on 11/11/80 and in 8 replicates from 4 sites on 11/19/80. In MW no coliforms were found in eight replicates from 14 sites on 11/19/80 and in BG no coliforms were found in four replicates from 7 sites on 12/2/80. Some of the replicate samples referred to in this paragraph were 200ml samples but since coliforms were not found, it seems reasonable to assume that coliforms would not have been found if only 100ml of those samples had been filtered. The results of all these sampling efforts tend to confirm the conclusion that several replicate samples per site may not reveal any coliforms.

One sample with no coliforms from a particular site is usually interpreted as a coliform density of less than 1/100ml at that site. Do several replicates from a particular site provide any more information? This question may be treated statistically using the binomial distribution. One sample provides only about 63% confidence that the coliform density at that site is actually <1/100ml, 2 samples about 85% confidence, 3 samples about 95% confidence, 4 samples about 98% confidence and so forth. On the other hand, one negative sample provides more than 99% confidence that the coliform density at the site is <10/100ml. Additional replicates do provide additional information, but the amount of information gained with each additional replicate decreases as the number of replicates increases.

5.6.5 Positive Sampling Results

It was already shown in Table 5.24 that no significant differences in the average coliform counts were found between the first and second 100ml replicates. Thus, although collecting and analyzing a second 100ml replicate may change the average coliform count, the change is not significant within the range of variability of the counts. Indeed, it would be very strange if the results did show a significant difference.

Does collecting second 100ml replicates increase the number of sites where coliforms are found? Comparison of columns 2 and 5 of Table 5.25 shows clearly that it does. The increase in the fraction of sites positive gained by taking the second replicate sample varies between 0.002 and 0.096. However, as might be expected, the increase in fraction positive is greatest for the systems where many coliforms were found and least where coliforms were rare. Thus, the gain in information about the overall level of contamination in the system is small. The values in column 3 of Table 5.25 are the binomial probabilities, $P=1-(1-p)^2$, for having either one or both of two replicates positive where p is the probability of a single sample having coliforms (column 2). The fact that the actual values in column 4 are less than the calculated values in column 3 merely shows that the two replicates are not independent measurements on the system.

Does collecting the second replicate help in locating sites where coliform densities are high (>4 per 100ml)? The last two columns of Table 5.25 shows that it sometimes helps, but not a great deal. The systems for which the greatest increase was found were those with very high levels of coliforms which were frequently found with only one sample.

All-in-all, collecting and analyzing a second 100ml replicate does not add a great deal of information. Systems with low levels of contaminations and systems with high levels of contamination show the same result with either one or two replicates per site if enough sites are sampled.

5.6.6 Larger Volume Samples

Interference with the detection of coliforms in 1000ml samples due to suspended matter was observed on several occasions. Thus, samples of that large a volume using the MF technique must be considered to be unreliable unless the interference problem can be overcome. This could be accomplished by filtering the sample in several portions. Actually, the same type of interference was observed in some 200ml samples collected in SR on 11/5/81. However, those were samples collected soon after the system had been flush by opening fire hydrants and were very turbid. Such interference is usually not seen in 200ml samples.

A comparison of sampling results obtained from both 100ml samples and 200ml is made in table 5.26. Data from days on which no coliforms were found were eliminated. Of course, only data from those days on which both 100ml

TABLE 5.24

ANALYSIS OF DATA FROM DUPLICATE 100ml SAMPLES

System Sampled	Nos. of Sites	First Replicate Fract. Pos.	95% Conf. Interv.	Mean Density (per 100ml)	95% Conf. Interv.	Second Replicate Fract. Pos.	95% Conf. Interv.	Mean Density (per 100ml)	95% Conf. Interv.	Both Replicates Fract. of Sites Positive	Mean Density (per 200ml)
Coatesville	185	.027	±.023	0.24	± 0.35	.011	±.015	0.02	± 0.04	.038	0.27
Woodbury Heights	215	.186	±.052	0.89	± 0.54	.181	±.051	1.48	± 0.80	.239	2.31
Brant Beach	177	.186	±.057	>4.67	± >2.46	.220	±.061	>6.37	± >3.08	.299	>8.37
Terrace Plant	255	.239	±.052	>3.21	± >1.61	.259	±.054	>4.68	± >2.43	.380	>6.89
Downingtown	174	.029	±.025	>0.50	± >0.90	.023	±.022	>0.10	± >0.14	.046	>0.60
Brooklawn	239	.017	±.016	>0.36	± >0.66	.025	±.020	>0.38	± >0.66	.033	>0.41
Mt. Idy	204	.010	±.014	>0.40	± >0.77	.015	±.017	>0.40	± >0.77	.015	>0.40
Spring Run	186	.065	±.035	>0.93	± >0.98	.086	±.040	>0.80	± >0.87	.124	>1.73
Marshallton Woods	98	.111	±.062	0.99	± 1.28	.051	±.043	0.18	± 0.26	.131	1.17
Bradford Glen	46	.109	±.089	>2.15	± >3.42	.087	±.081	>1.82	± >3.41	.152	>3.89

TABLE 5.25

EFFECT OF SECOND REPLICATE 100ml SAMPLE ON FRACTION OF SITES POSITIVE AND PERCENT OF SITES WITH 4 OR MORE COLIFORMS PER 100ml

System Sampled	Fraction of 100ml Samples Positive	Fraction of Sites Positive from 2-100ml Samples		Gain in Fraction of Sites Positive	Percent of Sites with ≥4 per 100ml	
		Calculated	Actual		First Replicate	Both Replicate
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coatesville	0.019	0.038	0.038	0.019	1.08	1.08
Woodbury Heights	0.184	0.0333	0.239	0.055	5.58	7.44
Brant Beach	0.203	0.365	0.299	0.096	10.73	14.12
Terrace Plant	0.249	0.436	0.380	0.131	10.58	15.29
Downingtown	0.026	0.051	0.046	0.020	1.15	1.72
Brooklawn	0.021	0.042	0.033	0.011	0.84	0.84
Mt. Idy	0.013	0.025	0.015	0.002	0.49	0.49
Spring Run	0.0755	0.145	0.124	0.048	3.76	4.30
Marshallton Woods	0.081	0.155	0.131	0.050	2.04	3.06
Bradford Glen	0.098	0.186	0.152	0.054	8.70	8.70

TABLE 5.26

COMPARISON OF 100ml AND 200ml SAMPLES

System Sampled (1)	Fraction of 100ml Samples Positive (2)	Fraction of Sites Positive from 2-100ml Samples (3)	Fraction of Sites Positive from 1 200ml Sample		Percent of Samp. with $\geq$ 4 Coliforms/100ml	
			First Replicate (4)	Second Replicate (5)	100ml Samples (6)	200ml Samples (7)
Downingtown	0.018	0.036	0.107	0	1.1	0.6
Mt. Idy	0.029	0.034	0.057	0.046	1.2	1.2
Spring Run	0.140	0.220	0.170	0.170	8.2	8.2
Marshallton Woods	0.099	0.185	0.072	0.074	2.4	1.9
Bradford Glen	0.265	0.412	0.294	0.294	11.8	23.5

Note: Data only from sampling days on which at least 2-100ml and 2-200ml replicates were taken at some sites and coliforms were found in all samples.

and 200ml replicates were taken could be used. Data from only the first two 100ml replicates and the first two 200ml replicates were used because there were only a few days on which additional replicate samples were collected.

The fraction of 200ml samples in which coliforms were found was usually greater than the fraction of 100ml samples in which coliforms were found (columns 4 and 5 versus column 2) but not much greater. The data from SR, MW, and BG suggest that two 100ml samples give a better chance of detecting coliforms at a sampling site than one 200ml sample (columns 4 and 5 versus column 3) but those differences are not statistically significant. In four of the five systems sampled the percent of samples with >4 coliforms per 100ml in 200ml samples was equal to or less than the percent for 100ml samples. Only in BG where only 17 sets of samples were used for the data in Table 5.26 were the 200ml samples better for detecting high coliform densities.

During the sampling of MW several replicate 100ml and 200ml samples were collected at each site sampled on three different days. It was expected that these sets of samples might show that 200ml samples were significantly better for detecting coliforms or high densities of coliforms than 100ml samples. However, the data do not demonstrate this because the fraction of samples positive and the coliform densities were actually greater for 100ml samples than for 200ml samples. This puzzling result is probably caused by the highly non-uniform dispersion of coliforms in these systems and should not be extrapolated.

The data used for comparison of 100ml and 200ml samples demonstrates mainly that coliform counts on samples from water distribution systems are quite variable. There is no empirical evidence to show that 200ml samples give a significantly greater probability of detecting coliforms or high densities (>/100ml) of coliforms than 100ml samples. Intuitively, it seems that 200ml samples should be better than 100ml samples. However, this intuition is explored from a theoretical approach in the next part of this section.

The following is intended to demonstrate that there is at least one theoretical model of coliform dispersion in water distribution systems which is consistent with the conclusions about 200ml sampling versus 100ml samples; that is, that they are not significantly different for detecting coliforms or high densities of coliforms.

Assume that the dispersion of coliform bacteria in some water distribution systems follows a lognormal distribution. Consider six such systems designated A, B, C, D, E, and F with geometric means and geometric standard deviations of 10^{-1} and 10, 10^{-3} and 100, 10^{-5} and 1000, 10^{-2} and 10, 10^{-4} and 100, and 10^{-6} and 1000 respectively. The values of the parameters of the distributions have been selected to conform to the ranges of these parameters estimated from sampling data. Systems A, B, and C have arithmetic mean coliform densities of more than one per 100ml and should fail to meet the microbiological MCL of the Regulations. Systems D, E, and F have arithmetic mean coliform densities of less than 1 per 100ml and should meet the microbiological MCL of the Regulations.

Assume that in some time period in which the bacteriological quality of the water does not change, 1000 samples are obtained from each of these six systems. The expected results from such samplings are summarized in Tables 5.27 and 5.28. The values of the fraction of sites sampled in each density range were obtained from a table of the standardized normal distribution using the logarithms of the densities to calculate the standardized normal variate, $Z = (X)/S$, from a statistical table. The numbers of samples expected from each density range were calculated by multiplying the fraction of sites by the assumed number of samples. The values in the last three columns were obtained by multiplying the number of samples in each density range by an assumed count in that density range and dividing by 1000. The samples with more than eighty coliforms per 100ml were averaged in as 80, those in the range of 10-80/100ml as 30, those in the range of 1-10/100ml as 3, those in the range of 0.1-1/100ml as 0.3, and so forth. Thus, the 21 samples collected from sites in system A where the coliform density is between 10 and 80 are assumed to have an average of 30 coliforms per sample for a total of 630 coliforms which contribute 0.63 to the arithmetic mean of the coliform count from 1000 samples. Notice that if enough samples are collected from sites where the coliform density is low, it is expected that some coliforms will turn up. For instance, the 191 samples collected from sites in system B where the coliform density is between 0.001 and 0.01 are expected to turn up one coliform. However, samples from sites where the coliform density is less than 10^{-3}/100ml are not expected to turn up coliforms because it would require 1000 standard 100ml samples of water with a density of 0.001 per 100ml to turn up one coliform. The significant item to notice from these tables is that the largest contributions to the arithmetic mean coliform densities are made by the few samples from sites with coliform densities of more than 10 per 100ml and for systems B, C, E, and F by the samples from sites where the density is more than 80 per 100ml.

The calculated values for the arithmetic mean, $\alpha = \exp [\ln X_g + 1/2(\ln S_g^2)]$ for the six distribution systems are: A, $\alpha = 1.417$; B, $\alpha = 40.29$; C, $\alpha = 229{,}949.74$; D, $\alpha = 0.142$; E, $\alpha = 4.029$; and F, $\alpha = 22{,}994.97$. The large differences between the theoretical arithmetic means and those calculated for a hypothetical set of 1000 samples is due to the occurrence of samples with more than 80 coliforms per 100ml. The use of the lognormal distribution to represent the frequencies of various counts forces the conclusion that some of the TNTC samples contain many more than eighty coliforms per 100ml. One sample with thousands or millions of coliforms would completely dominate the theoretical arithmetic mean but was averaged in as >80/100ml for the hypothetical sample sets.

In this analysis, it is assumed that collection of two 100ml samples at a given site is equivalent to collecting one 200ml sample at the same site. A consequence of this assumption is that, on the average, 200ml samples will turn up twice as many coliforms as 100ml samples. This assumption is not inconsistent with the sampling data presented earlier; although the data do not demonstrate it.

Table 5.29 gives the expected numbers of counts of 0, 1, 2, and 3 coliforms per sample for both 100ml and 200ml samples from sites with densities in the range of 0.001 to 10 per 100ml. These counts were obtained by assum-

TABLE 5.27

EXPECTED RESULTS OF SAMPLING FROM LOGNORMAL DISTRIBUTIONS

(Systems with More Than 1 Coliform per 100ml)

Range of Counts (per 100ml)	Fraction of Sites with Coliform Counts in Indicated Range			Expected Number of Samples in Indicated Range out of 1000 Samples			Contribution to Arithmetic Mean of 1000 Samples		
	A	B	C	A	B	C	A	B	C
>80 (TNTC)	0.0019	0.0071	0.0107	2	7	11	0.160	0.560	0.880
10-80	0.0209	0.0157	0.0121	21	16	12	0.630	0.480	0.360
1-10	0.1359	0.0440	0.0247	136	44	25	0.408	0.132	0.075
0.1-1	0.3413	0.0919	0.0443	341	92	44	0.102	0.028	0.013
$10^{-2}-10^{-1}$	0.3413	0.1498	0.0669	341	150	67	0.010	0.005	0.002
$10^{-3}-10^{-2}$	0.1359	0.1915	0.0927	136	191	93	-	0.001	-
$10^{-4}-10^{-3}$	0.0215	0.1915	0.1193	22	191	119	-	-	-
$<10^{-4}$	0.0013	0.2417	0.6293	1	309	629	-	-	-
						Arithmetic Means	1.310	1.206	1.330

System A: $\overline{X}_g = 0.1$, $S_g = 10$; System B: $\overline{X}_g = 10^{-3}$, $S_g = 100$; System C: $\overline{X}_g = 10^{-5}$, $S_g = 1000$

TABLE 5.28

EXPECTED NUMBERS OF COUNTS IN SAMPLES FROM MIDDLE RANGE COLIFORM DENSITY SITES

(Based on 1000 Total Samples)

System	Density Range (per 100ml)	No. of 100ml Samples with Counts of 0	1	2	3	Total Samples	Total Counts	No. of 200ml Samples with Counts of 0	1	2	3	Total Samples	Total Counts
A	1-10	7	20	30	30	136	408	0	2	6	12	136	816
	0.1-1	252	77	11	1	341	102	187	112	34	8	341	204
	0.01-0.1	331	10	0	0	341	10	321	20	0	0	341	20
	0.001-0.01	136	0	0	0	136	0	135	1	0	0	136	1
B	1-10	2	7	10	10	44	132	0	1	2	4	44	264
	0.1-1	68	20	4	0	92	28	50	30	10	2	92	56
	0.01-0.1	145	5	0	0	150	5	140	10	0	0	150	10
	0.001-0.01	190	1	0	0	191	1	190	1	0	0	191	1
C	1-10	1	4	6	6	25	75	0	0	1	2	25	150
	0.01-1	32	11	1	0	44	13	24	15	4	1	44	26
	0.001-0.01	65	2	0	0	67	2	63	4	0	0	67	4
	0.001-0.01	93	0	0	0	93	0	92	1	0	0	93	1
D	1-10	1	3	5	5	22	66	0	0	1	2	22	132
	0.1-1	100	31	5	0	136	41	74	45	14	3	136	82
	0.01-0.1	331	10	0	0	341	10	321	20	0	0	341	20
	0.001-0.01	340	1	0	0	341	1	339	2	0	0	341	2
E	1-10	1	2	4	4	16	48	0	0	1	1	16	96
	0.01-1	32	11	1	0	44	13	24	15	4	1	44	26
	0.01-0.1	89	3	0	0	92	3	86	6	0	0	92	6
	0.001-0.01	150	0	0	0	150	0	149	1	0	0	150	1
F	1-10	1	2	3	3	13	39	0	0	0	1	13	78
	0.1-1	19	5	1	0	25	7	14	8	3	0	25	14
	0.01-0.1	43	1	0	0	44	1	42	2	0	0	44	2

TABLE 5.29

EXPECTED RESULTS OF SAMPLING FROM LOGNORMAL DISTRIBUTIONS

(Systems with Less Than 1 Coliform per 100ml)

Range of Counts (per 100ml)	Fraction of Sites with Coliform Counts in Indicated Range			Expected Number of Samples in Indicated Range out of 1000 Samples			Contribution to Arithmetic Mean of 1000 Samples		
	D	E	F	D	E	F	D	E	F
>80	<.001	.0016	.0043	–	2	4	–	0.160	0.320
10–80	.0013	.0046	.0056	1	5	6	.030	0.150	0.180
1–10	.0215	.0164	.0129	22	16	13	.066	0.048	0.039
0.1–1	.1359	.0440	.0247	136	44	25	.041	0.013	0.007
10^{-2}–10^{-1}	.3413	.0919	.0443	341	92	44	.010	0.003	0.001
10^{-3}–10^{-2}	.3413	.1498	.0669	341	150	67	.001	–	–
10^{-4}–10^{-3}	.1359	.1915	.0927	136	191	93	–	–	–
$<10^{-4}$	.0228	.4332	.7486	23	500	748	–	–	–
							0.138	0.374	0.547

System D: $\overline{X}g = 10^{-2}$, Sg = 10; System E: $Xg = 10^{-4}$, Sg = 100; System F: $\overline{X}g = 10^{-6}$, Sg = 1000 ing that the coliforms from samples in a given density range are randomly distributed among those samples and calculation of the counts from Poisson probabilities. For instance, the 341 samples from sites in system A where the coliform density is between 0.1 and 1 per 100ml are expected to yield 102 coliforms for an average of 0.299 per 100ml. The probability of obtaining a count of zero (negative sample) is $e^{-0.299}$ or 0.7415, which multiplied by 341 samples gives 252.9 samples. The 252.9 samples would usually round off to 253 but the most likely way to distribute the 102 coliforms is in 77 samples with one coliform, 11 samples with 2, and 1 sample with 3. This leaves 252 samples with no coliforms. It is assumed that if those samples were 200ml samples (or two 100ml samples) they would yield 204 coliforms for an average of 0.598 per 200ml. The probability of obtaining a count of zero in this case is $e^{-0.598}$ or 0.550 which when multiplied by 341 samples gives 187 samples.

This method of distributing the coliforms among samples could be called a segmented lognormal-Poisson model. It is necessary to change, at some point, from the lognormal which is a continuous distribution to a discrete distribution to obtain integral counts per sample. The Poisson is the simplest discrete distribution to use for this purpose and the parameters of the lognormal distribution are preserved. The division of the density ranges into factors of 10 is arbitrary. They could be divided into smaller ranges but finer division of the density ranges would not significantly change the results of these calculations.

It is expected that all samples collected from sites where the density is greater than 10 per 100ml will have some coliforms present and that all samples collected from sites where the density range is less than 0.001 per 100ml will have no coliform present. Thus, the fraction of the sites where coliforms are found will be changed by collecting two 100ml samples or one 200ml sample per site only where the density is in the range between 0.001 and 10 per 100ml. This is illustrated in Tables 5.30 and 5.31. The fraction of the sites sampled changes only a small amount when the sample volume is doubled because only a portion of the sites have a coliform density in the range of 0.001 to 10 per 100ml.

This result is not intuitively obvious. It can be made clearer by dividing the sample sites into three categories. At sites where the coliform density is greater than 10 per 100ml, there is an extremely high probability that some coliforms will be found in either 100ml or 200ml samples. At sites where the density is much less than 1 per 100ml, the probability of finding a coliform is twice as high for 200ml but still so low that many samples must be collected before even one coliform is found. (Note that the 1000 samples used in the example is far more than is required for any distribution system in a months time). The sample volume will make a difference only where the coliform density is close to 1 per 100ml.

The sample volume should have no effect on the measured coliform density. This was found, within the sampling variability, for the systems in which both 100ml and 200ml replicates were collected. Doubling the sample volume will only slightly increase the number of sites where coliforms are found. This is also true within the variability of the sampling results for those

TABLE 5.30

EFFECT OF SAMPLE VOLUME ON FRACTION OF SITES WITH COLIFORMS
(Systems with More Than 1 Coliform per 100ml)

Density Range (per 100ml)	100ml Samples A No.of Pos. Sites	100ml Samples A No.of Neg. Sites	100ml Samples B No.of Pos. Sites	100ml Samples B No.of Neg. Sites	100ml Samples C No.of Pos. Sites	100ml Samples C No.of Neg. Sites	200ml Samples A No.of Pos. Sites	200ml Samples A No.of Neg. Sites	200ml Samples B No.of Pos. Sites	200ml Samples B No.of Neg. Sites	200ml Samples C No.of Pos. Sites	200ml Samples C No.of Neg. Sites
>80	2	0	7	0	11	0	2	0	7	0	11	0
10-80	21	0	16	0	12	0	21	0	16	0	12	0
1-10	129	7	42	2	24	1	136	0	44	0	25	0
0.1-1	89	252	24	68	12	32	154	187	42	50	20	24
0.01-0.1	10	331	5	145	2	65	20	321	10	140	4	63
0.001-0.01	0	136	1	190	0	93	1	135	1	190	1	96
10^{-4}--10^{-3}	0	22	0	191	0	119	0	22	0	191	0	119
$<10^{-4}$	0	1	0	309	0	629	0	1	0	309	0	629
Total	251	749	95	905	61	939	334	666	120	880	73	927
Fraction Positive	0.251		0.095		0.061		0.334		0.120		0.073	

TABLE 5.31

EFFECT OF SAMPLE VOLUME ON FRACTION OF SITES WITH COLIFORMS

(Systems with Less Than 1 Coliform per 100ml)

	100 ml Samples						200 ml Samples					
	D		E		F		D		E		F	
Density Range	No.of Pos. Sites	No.of Neg. Sites	No.of Pos. Sites	No.of Neg. Sites	No.of Pos. Sites	No.of Neg. Sites	No.of Pos. Sites	No.of Neg. Sites	No.of Pos. Sites	No.of Neg. Sites	No.of Pos. Sites	No.of Neg. Sites
>80			2	0	4	0	-		2	0	4	0
10-80	1	0	5	0	6	0	1	0	5	0	6	-
1-10	21	1	15	1	12	1	22	0	16	0	13	0
0.1-1	36	100	12	32	6	19	62	74	20	24	11	14
0.01-0.1	10	331	3	89	1	43	20	321	6	86	2	42
0.001-0.01	1	340	0	150	0	67	2	339	1	149	0	67
10^{-4}-10^{-3}	0	136	0	191	0	93	0	136	0	191	0	93
$<10^{-4}$	—	23	—	—	—	—	—	—	—	—	—	—
	69	961	37	963	29	971	107	893	50	950	36	964
Fraction Positive	0.069		0.037		0.029		0.107		0.050		0.036	

systems where 100ml and 200ml samples were collected. In terms of a monitoring program, the information gained either by taking two 100ml replicates or by doubling the volume of the standard sample is very small. On the other hand collecting and analyzing 200ml replicates is no more expensive than collecting and analyzing 100ml replicates if the MF technique is used for examination and the samples are transported by vehicle. Sending the larger volume samples by mail would cause some extra expense. An increase in the frequency of occurence of interference due to particulate matter on the membrane filter could be expected with larger volume samples. Using larger volume samples could have some advantages for larger systems which have their own microbiological laboratories and can deal with the interference problem. However, for smaller systems, it could be a disadvantage without equivalent benefit.

SECTION 6

MECHANISTIC APPROACH

6.1 INTRODUCTION

The mechanistic approach is an effort to make use of physical information about the system such as the sizes and locations of the pipes, the layout of the pipes, the amount of water in the system, and the flow of water through the system as a guide for microbiological monitoring of a water distribution system. Various components of the mechanistic approach are explored in this section. A method of describing the configuration or layout of the distribution system is developed. Then several hypotheses about mechanisms by which coliform bacteria are dispersed in a water distribution system are examined. The goal of this approach was to develop information which could be useful in identifying areas where there is a high probability of coliform contamination.

6.2 CONFIGURATION OF SECTIONS

A water distribution system can be divided into isolated sections based on projected water flow patterns. This division aids in the design of sampling programs. By recognizing sections of a system, samples may be taken from each section and hence each segment of the water flow. Sections delineated on the basis of pipe layout and hence water flow patterns may have different probabilities of coliform occurrence. Recognition of this could aid in understanding why coliforms are or are not found within different parts of a system.

6.2.1 Definitions of Distribution System Elements

According to the National Interim Primary Drinking Water Regulations (EPA 1977) a water system means a system for the provision of piped water for human consumption and includes collection, treatment, storage, and distribution facilities. We consider the distribution system to include the pipes used to convey water after collection and treatment (if any) to the service connections plus the storage facilities connected directly to such pipes. In some cases, although this did not occur in any of the systems we sampled, chlorine may be added to the water already in the distribution system particularly as it is withdrawn from storage.

We have designated the water conveyance elements (pipes) of a water distribution system as follows: A transmission main is a pipe which carries water from the source (well or treatment plant) to the distribution system. Transmission mains are typically the largest pipes of the system. Any bacteriological quality changes which occur in the transmission mains may affect water quality in the entire system. A distribution main is a pipe which carries water into or out of an isolated section of the system. It has a distinct transmission function but the area served is more limited than that of the transmission main. Flow reversals in a distribution main occur only when it is connected with a storage tank.

A street lateral is a pipe which serves to deliver water to customer services within a single city block and sometimes transmits water to other street laterals. We call a street lateral which transmits water to at least two other street laterals a transmission street lateral. Flow reversal in transmission street laterals may occur frequently due to changes in water demand. A street lateral which supplies water only for local use is called a peripheral street lateral, of which there are two types. A street lateral which delivers water only to customer services within a single city block and usually ends in either a fire hydrant or a flush out valve is called a dead end. Flow reversal is impossible and the water is semi-stagnant. A loop consists of two dead end street laterals connected together.

In addition to the water conveyance elements of a distribution system, we define pumping stations and storage facilities. Pumping Stations are located on transmission mains or on distribution mains and are intended to increase the water pressure on the downstream side. They may be used as sampling locations for determination of the water quality supplied to a section or several sections of the distribution system. Distribution system storage is either elevated or is operated in conjunction with a pumping station. The purpose is to maintain pressure in parts of the system and to serve as a reservoir of water to meet extraordinary demand such as fire flow. Standpipes and storage tanks usually have large volumes of water which mix and interchange only slowly with water which is actually distributed to services. Uncovered reservoirs are well known to be sources of bacterial contamination and have been eliminated from many distribution systems. Even covered reservoirs, storage tanks, and standpipes may allow access of birds, rodents, plant materials, dust, etc. to the water in the distribution system and are likely sites of bacterial contamination of water in the system.

6.2.2 Isolated Sections

An isolated section of a water distribution system is a collection of street laterals and distribution mains which supply water for local use to several blocks. There is input to the isolated section from at least one but no more than two mains. There may be output from the section to other isolated sections. Water quality changes which occur within an isolated section of a distribution system may not be measured by samples collected from other parts of the distribution system. Therefore, some samples must be collected in each isolated section to insure that the samples as a whole are representative of water quality conditions throughout the entire system.

The four configurations which the pipes that make up an isolated section can have are linear, loop, dendritic, and grid, as illustrated in figure 6.1. An isolated section can have a compound configuration; i.e., a combination of two of the basic patterns; but not a complex configuration; i.e., more than

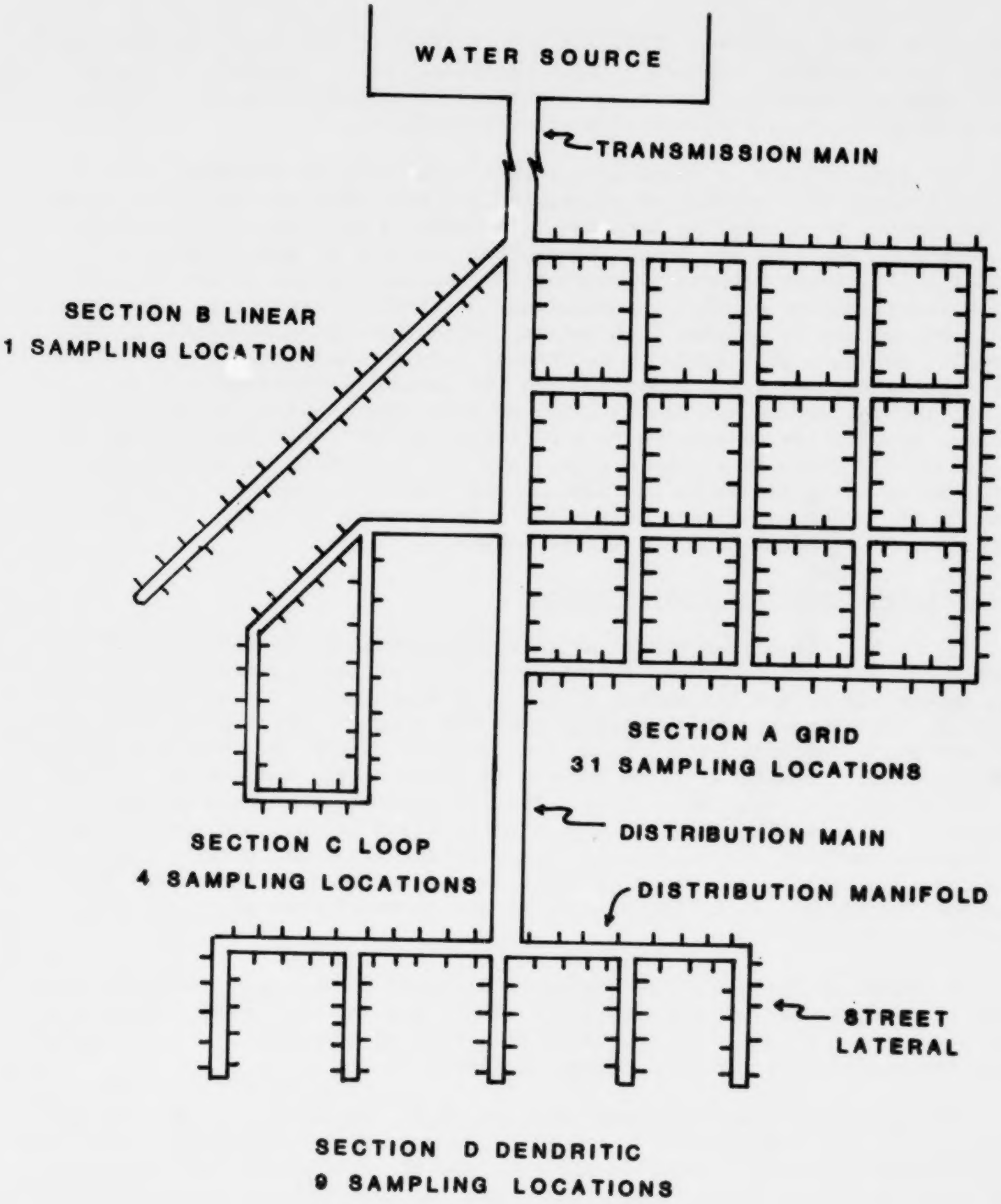

WATER SOURCE
TRANSMISSION MAIN
SECTION B LINEAR
1 SAMPLING LOCATION
SECTION A GRID
31 SAMPLING LOCATIONS
DISTRIBUTION MAIN
SECTION C LOOP
4 SAMPLING LOCATIONS
DISTRIBUTION MANIFOLD
STREET
LATERAL
SECTION D DENDRITIC
9 SAMPLING LOCATIONS

Figure 6.1 Representative Examples of Four Isolated Section Configurations.

two of the basic patterns. If a mixture of three of the basic patterns is found, it is divided into two or more sections. Thus, in addition to the four basic configurations, there are six possible compound configurations for a total of ten possible configurations overall.

For this project, a block on a street which usually corresponds to a street lateral or a portion of a distribution main with service connections is designated as a sampling location. The ends of the pipes corresponding to sampling locations are joints connecting with two or more other pipes. Usually, the street lateral is underneath a street and the joints are underneath intersections so that the sampling location corresponds to a block on a street and may be sampled from service connections on either side of the street. However, this needs to be checked from a blueprint of the distribution system because there sometimes are two laterals underneath a street, a street lateral may be two blocks long, or some other unusual situation may occur. Any service connection in a building, either residential, commercial, or publically owned is a sample site. Any tap or hydrant in or associated with the building should be the same as any other for sampling purposes as long as the water is allowed to run for a few minutes to clear the service line and obtain water from the street lateral.

6.2.3 Configuration of Systems Sampled

Table 6.1 summarizes the structural parameters of the nine systems sampled. The number of sampling locations is more or less proportional to the population served but the number of isolated sections is a function of other factors. The largest system sampled (LB) was a single large grid which was divided into four sections based on the location of the two treatment plants which provide the water. It is located on flat terrain and was not crossed by any highways or railroad tracks. The two systems which have the most isolated sections (CV and DT) are located in hilly terrain and are divided by highways and railroads. WH and BL are relatively small systems which are divided into isolated sections by highways and railroads. The other four systems are located on hilly terrain but are so small that they have few isolated sections.

As shown in Table 6.1 the number of sampling locations corresponds very well with the number of pipes (one per block) and the number of joints (one for every two blocks). The number of dead ends is more variable and depends upon the configuration of individual sections.

The last four systems sampled are too small, to contribute much to the analysis by isolated sections. The first five systems sampled had a total of twenty eight sections. Eight of the ten possible configurations were found. The most frequently found configuration was the dendritic grid(10) followed by the grid(6). The other configurations found were dendritic(4), linear loop(3), loop grid(2), linear(1), linear grid(1), and dendritic loop(1). Thus, the grid and grid combinations comprised 68% of the sections delineated. The dendritic pattern and its combinations comprised 54%; the loop and its combination, 21%; and the linear and its combinations, 18%. The linear-loop and loop-dendritic were not found in any of these five systems but were found in the very small systems.

TABLE 6.1

STRUCTURAL ANALYSIS OF SYSTEMS SAMPLED

System	Population Served	No. of Isolated Sections	No. of Sampling Locations	No.of Pipes	No. of Joints	No. of Dead Ends	Joints per Location	Dead Ends per Location
CV	20,000	9	453	349	196	58	0.43	0.12
WH	3,600	5	125	127	63	27	0.50	0.22
LB	3,000 45,000	4	644	635	319	137	0.50	0.21
DT	9,800	7	184	170	106	17	0.58	0.09
BL	2,700	3	74	73	40	10	0.54	0.14
MI	125	1	11	11	7	5	0.64	0.54
SR	550	3	17	18	8	4	0.47	0.23
MW	100	1	2	2	2	0	1	0
BG		1						

Table 6.2 summarizes information on the structural parameters of the five configurations for which more than one example was found. The size of an isolated section is measured by the number of sampling locations in that section. The structural parameters for the section show how well the physical make up of the distribution system corresponds with the layout of city blocks for that section. The ratios of pipes and joints per sampling location found for the entire systems also held for the individual sections. The number of dead ends per sampling location did vary with configurations. Dendritic configurations have an abundance of dead ends because that is basic to the definition. Grids have the smallest portions of dead ends. This appears to be due to the large number of joints and pipes which made up a grid and the fact that if many dead ends are found, the section was designated as a dendritic grid. The most dead ends per location, joint, or pipe are in dendritic configurations. The other common configurations have varying numbers of dead ends per location, joint, or pipe. These relative numbers of dead ends are large for these sections but they are not given a dendritic designation as the other properties of the system dominated the configuration.

TABLE 6.2

SUMMARY OF STRUCTURAL PARAMETERS BY SECTION CONFIGURATION

Configuration	Grid	Dendritic Grid	Loop Grid	Dendritic	Loop Linear
Number of Examples	6	10	2	4	3
Pipes per Location	0.89	0.93	0.86	1.0	0.93
Joints per Location	0.51	0.48	0.57	0.54	0.65
Dead Ends per Location	0.07	0.20	0.11	0.36	0.25
Dead Ends per Pipe	0.08	0.23	0.20	0.31	0.28
Dead Ends per Joint	0.14	0.46	0.21	0.65	0.38
Pipes per Joint	1.69	1.97	1.57	2.06	1.40

These analyses demonstrate that choosing sample locations on the basis of street patterns often parallels the more complicated procedure of choosing sample locations on the basis of the layout of the actual distribution mains. There are descrepancies however. In one section of System CV the number of lines per location and joints per location are exceptionally low (0.44 and 0.18, respectively). This is because of a dendritic pattern in which street laterals extend through several blocks without joints to other street laterals. The question that arises is, given a limited number of samples, might it be more efficient in such an instance, to make a sampling location one street lateral rather than one block?

The parameters of greatest variability are those involved with the number of dead ends. As dead ends are associated with decreased water flow, sections containing a large number of dead ends were postulated to have greater probability of coliforms from regrowth than other sections (Pipes and Christian 1978, 1980).

Other aspects of section structure include the pipe length, volume and average pipe cross sectional area. These parameters were calculated and are presented in table 6.3. Pipe diameter and length for each location were cataloged, the total length of each pipe size was summed for each isolated section, and volume and average cross section of mains for the various sections. No one section configuration represented the greatest length or volume of pipe. The smallest sections by either criterion were linear loop and linear. Sizes of the other section types were medium to large with no obvious pattern. If the flow pattern of water does in fact relate to section characteristics, it is expected that the volume of the section may be positively correlated with the cross sectional area. Larger sections require larger mains on average for the maintenance of adequate water supply. A scatter diagram of the data is presented in figure 6.2. To examine this hypothesis, a Kendall's rank correlation analysis and a Spearman's rank correlation analysis were conducted between section volume and crosssectional area. Both demonstrated a significant correlation for 28 sections analyzed. The Kendall's rank coefficient of correlation was 0.360 providing a $t_s[H_0T=0]$ of 2.74 with a probability that the t_s arose by chance of 0.0062. The Spearman's rank coefficient of correlation was 0.44 with a probability that the correlation occured by chance of less than 0.02. The major contributing factor to the significance of correlation appears to be the presence of large transmission mains in the large sections as would be expected.

6.2.4 Relationship of Configuration to Sampling

Clearly, samples should be collected from all of the isolated sections of a water distribution system because contamination occurring in one section usually does not affect water quality in other sections. However, it is not clear that the number of samples collected and the locations used within an isolated section should depend upon the configuration of that section. The direction of the flow of water can be established in a rather straightforward manner for linear and dendritic sections, and individual sampling locations can be designated as either upstream or downstream of other locations. On the other hand, the direction of flow in individual lines in a grid or loop configuration is not obvious and probably changes from time to time with

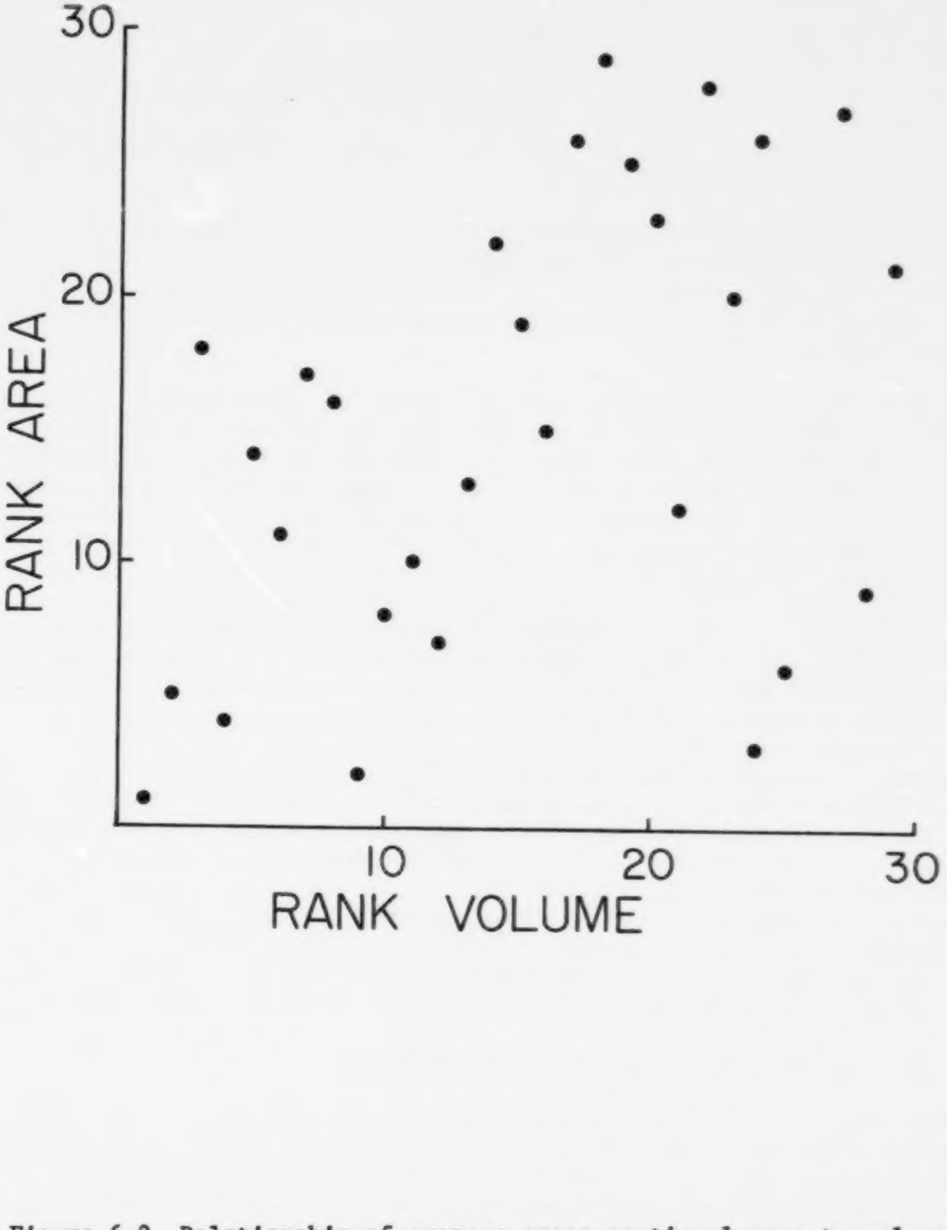


Figure 6.2 Relationship of average cross sectional area to volume of water distribution system sections.

TABLE 6.3

DIMENSIONS OF DISTRIBUTION SYSTEM PIPES BY SECTION

System	Section	Configuration	Average Pipe Area(ft^2)	Total Pipe Length(ft)	Volume of Water	
					(ft^3)	(gals)
CV	West	Grid	0.399	31,609	12,614	94,359
	Central	"	0.332	98,704	32,791	245,298
	N. Central	Dendritic	0.240	8,708	2,091	15,642
	E.F.	"	0.195	11,220	2,199	16,449
	Caln	Dend.-Grid	0.209	15,118	3,159	23,609
	S.C.	Dend.-Loop	0.229	21,355	4,881	36,513
	Foundry	Linear-Loop	0.280	8,220	2,300	17,205
	Tooth	" "	0.269	3,985	1,070	8,004
	North	Grid-Loop	0.179	6,880	1,229	9,194
	Total		0.307	202,589	62,331	466,273
WH	W. Central	Grid	0.174	15,500	2,698	20,182
	North	Dend.-Grid	0.193	16,750	3,240	24,232
	Southwest	" "	0.196	14,675	2,881	21,551
	East	Dendritic	0.369	18,400	6,784	50,753
	E. Central	"	0.203	8,100	1,650	12,345
	Total		0.235	73,425	17,253	129,063
LB	North	Grid	0.215	33,165	7,123	60,433
	N. Central	Dend.-Grid	0.185	56,636	10,485	78,533
	S. Central	" "	0.198	66,330	13,130	98,212
	South	" "	0.161	65,925	10,627	79,495
	Total		0.186	222,056	41,365	316,673
DT	Central	Grid	0.304	33,045	10,034	75,060
	North	"	0.459	15,735	7,230	54,077
	Southwest	Dend.-Grid	0.394	13,765	5,420	40,544
	East	" "	0.354	19,680	6,973	52,162
	West	Loop-Grid	0.363	32,205	11,690	87,447
	Northwest	Loop-Linear	0.210	7,300	1,535	11,483
	Southeast	Linear	0.569	10,040	5,712	42,729
	Total		0.369	131,770	48,595	363,516
BL	Central	Grid	0.350	9,965	3,492	26,122
	Southeast	Dend.-Grid	0.298	13,150	3,917	29,294
	Northeast	Dend.-Grid	0.237	9,970	2,364	17,676
	Total		0.295	33,085	9,771	73,092

changes in water demand. Thus, samples collected anywhere in a section within a grid or loop configuration may be considered to representative of the entire section whereas samples collected in a section with a linear or dendritic configuration may be used to isolate contamination in a street lateral or possibly even in a portion of a street lateral. These ideas are explored further by testing of specific hypotheses.

6.3 SAMPLING BY SECTIONS

Isolated sections of a distribution system are defined so that there is some hydraulic separation. Microbiological contamination occurring in one section may not be detected in samples collected from another section of the same distribution system. Of course, it is possible that contamination could be general throughout the system and would be detectable in samples collected in any part of the system. The lack of significant differences among sections of a distribution system would not prove that the concept of isolated sections does not hold but finding an instance of significant differences between two sections of the same system would establish the validity of the concept.

The other aspect of the question about the validity of the concept of isolated sections is how large can an isolated section be. In some water distribution systems, hundreds of city blocks may be supplied water from mains and street laterals which are part of the same large grid. If there is no hydraulic separation could contamination entering the distribution system be detected by samples collected 10 blocks away, or 20, 30, or 40 blocks away?

The data collected during this project are analyzed to determine what light they shed on these two questions. First, are there significant difference among isolated sections of the same distribution system? Second, in the absence of a clear hydraulic separation, do adjacent sampling sites or adjacent blocks give equivalent results?

6.3.1 Differences Among Sections

The coliform and standard plate count data for the six systems which had more than one isolated section are summarized in tables 6.4 through 6.9. The data for systems WH, LB, and BL are further subdivided by sampling period. These analyses are based on the results in the first two 100ml replicate samples collected from each site sampled in order to give an equivalence of data. The coliform counts and standard plate counts for each site sampled are the average of the counts obtained on the first two replicates.

Differences in the extent of microbiological contamination among the isolated sections of a distribution systems might be seen as statistically significant differences in the mean coliform densities, in the mean standard plate counts, or in the fraction of the samples in which coliform organisms are found.

Significant differences among the mean coliform counts or among the mean standard plate counts could be sought using analysis of variance. However,

TABLE 6.4

ANALYSIS OF SYSTEM CV BACTERIOLOGICAL WATER QUALITY PARAMETERS BY SECTION

Section	Number of Sites Sampled	Fraction with Coliforms	Coliform Density (per 100ml)		Standard Plate Count (per ml)	
			mean	var.	mean	var.
West	68	0.059	0.09	1.01	8.8	1600
North	24	0.083	0.06	0.05	0.7	4.10
North-Central	10	0	0	-	7.6	370
Central	83	0.036	0.14	1.37	7.6	3600
Caln	9	0	0		4.8	150
East				-		
Fallowfield	5	0	0	1.45	0.20	0.20
Foundry	11	0.091	0.36	-	1.0	10
Toth	5	0	0		0.47	0.76
South				-		
Coatesville	10	0	0		0.23	0.12

TABLE 6.5

ANALYSIS OF SYSTEM WH BACTERIOLOGICAL WATER QUALITY PARAMETERS BY SECTION

Section	Sampling Period	Number of Locations Sampled	Fraction with Coliforms	Coliform Density (per 100ml)		Standard Plate Count (per ml)	
				mean	var.	mean	var.
East	I	9	0	0	-	0.61	1.45
	II	23	0.304	0.74	2.56	20.02	3800
	III	23	0.087	0.26	0.95	6.20	120
East-Central	I	12	0.250	0.92	12.50	1.98	36.32
	II	16	0.312	0.50	2.19	3.92	34.53
	III	17	0.118	0.12	0.14	4.28	34.60
West-Central	I	6	0	0	-	10.96	630
	II	37	0.514	0.35	84.76	19.05	2600
	III	37	0.027	0.03	0.03	3.43	49.42
Southwest	I	17	0.176	1.47	13.13	11.56	580
	II	40	0.450	1.53	11.88	21.66	1900
	III	43	0.209	>3.32	>88.17	3.74	87.92
North	I	22	0	0	-	4.73	170
	II	38	0.316	0.54	1.26	36.0	5900
	III	53	0.246	>2.17	>66.08	18.54	4100

Sampling Periods
I = April-May 1979 (2 weeks)
II = May 1979 (2 weeks)
III = May-June 1981 (2 weeks)

TABLE 6.6

ANALYSIS OF SYSTEM LB BACTERIOLOGICAL WATER QUALITY PARAMETER BY SECTION

Section	Sampling Period	Number of Locations Sampled	Fraction with Coliforms	Coliform Density (per 100ml)		Standard Plate Count (per ml)	
				mean	var.	mean	var.
North	I	43	0.140	0.95	33.41	9.34	960
	II	35	0.371	>8.78	>52	>130	>100
North-Central	I	49	0.122	0.14	0.35	23.46	1100
	II	58	0.448	9.78	>39	13.91	1300
South-Central	I	62	0.274	>2.30	>49.56	12.54	1700
	II	63	0.317	>3.94	>180	>310	$>10^6$
South	I	61	0.328	>3.0	>71.03	>250	$>10^6$
	II	76	0.421	>7.95	>340	>260	$>10^6$

Sampling Periods I = June 1979 (2 weeks)
II = July 1979 (2 weeks)

TABLE 6.7

ANALYSIS OF SYSTEM DT BACTERIOLOGICAL WATER QUALITY PARAMETERS BY SECTION

Section	Number of Locations Sampled	Fraction with Coliforms	Coliform Density (per 100ml)		Standard Plate Count (per ml)	
			mean	var.	mean	var.
West	48	0.021	0.05	0.13	2.40	50.39
Northwest	12	0.083	0.042	0.02	2.98	38.78
Southwest	18	0.056	0.06	0.0	1.01	0.96
Central	39	0.051	0.19	0.86	2.13	69.24
East	21	0	0	-	0.75	1.06
North	27	0.148	0.09	0.06	3.41	200
Southeast	9	0	0	0	2.27	9.46

TABLE 6.8

ANALYSIS OF SYSTEM BL BACTERIOLOGICAL WATER QUALITY PARAMETERS BY SECTION

Section	Sampling Period	Number of Locations Sampled	Fraction with Coliforms	Coliform Density (per 100ml)		Standard Plate Count (per ml)	
				mean	var.	mean	var.
Southeast	I	81	0.049	>1.01	>19.50	4.89	1100
	II	50	0.080	0.110	0.46	11.29	960
Central	I	84	0.024	0.01	0.43	1.46	18.40
	II	56	0.143	0.18	0.51	4.05	50.24
Northeast	I	70	0.029	0.01	0.01	1.73	23.92
	II	64	0.109	1.06	8.43	17.21	2400

Sampling Periods I = March-April 1980
II = June 1981

TABLE 6.9

ANALYSIS OF SYSTEM SR BACTERIOLOGICAL WATER QUALITY PARAMETERS BY SECTION

Section	Sampling Period	Number of Locations Sampled	Fraction with Coliforms	Coliform Density (per 100ml)		Standard Plate Count (per ml)	
				mean	var.	mean	var.
East		79	0.101	>0.75	>20.80	10.15	1300
West		44	0.136	0.35	1.670	27.93	2600
North		53	0.151	0.91	15.37	1.90	12.04

inspection of tables 6.4 through 6.9 shows that in most cases the variances are much greater than the mean. The only cases in which the variances are low are when the samples were collected in the winter or when the number of samples collected in a section was relatively few. The high values for the variances indicate that seeking significant differences by analysis of variance on the untransformed data would be futile. A rational for using a data transformation to stabilize the variances could be developed. However, since most of the counts were zeroes the data would have to be coded and then transformed. This could introduce artifacts into the analysis of variance and any significant differences found would be questionable. Therefore, it was decided not to use this approach in searching for significant differences among sections.

If there is a significant difference in the proportion of the sites where coliforms are found among several sections, it can reasonably be concluded that there is a different level of microbiological contamination among the sections. Significant differences in proportions can be sought using a χ^2 contingency table analysis. Such analysis can be applied in a straightforward manner for most of the systems and most of the sampling periods. However, one of the requirements for the analysis is that the expected number of positive sites for each section, which is obtained by multiplying the overall fraction positive by the number of sites sampled in the section, must be greater than 1. This requirement is violated for system CV where the overall fraction positive was 0.044 which means that 23 sites would have to be sampled to give an expected number of positive sites of 1.02, for the first sampling period for system WH where the overall fraction positive was 0.091 and 11 sites would have to be sampled to give an expected number of positive sites of 1.001, and for system DT where the overall fraction positive was 0.052 and 20 sites would have to be sampled to give an expected number of positive sites of 1.04. The only way to meet this requirement for the χ^2 test is to group some sections together. When sections are grouped, it is necessary to group contiguous sections.

The results of the χ^2 contingency table analyses are presented in table 6.10. For system CV the North and North Central sections were grouped together, as were the Central and Caln sections, and the last four sections. For the first sampling period for system WH, the East, East Central, and West Central sections were grouped together. For system DT, the Northwest and Southwest sections were grouped together as were the North, East, and Southeast sections. The null hypothesis (H_0) for these analyses is that there are no differences among the sections of each system for each sampling period in the fraction of the sites where coliforms were found. The only instances where the null hypothesis can be rejected at the 5% level are the third sampling period for system WH and the first sampling period for system LB.

It is instructive to pursue the differences detected further. The LB system is one large interconnected grid which is supplied water from two different sources. The separation into isolated sections is based on the location of the treatment plants. The North Section is north of the Brant Beach Plant and the South Section is south of the Terrace Plant. The separation between the North Central and South Central Sections was made halfway

TABLE 6.10

CONTINGENCY TABLE ANALYSES FOR DIFFERENCES AMONG SECTIONS

System	Sampling Period	Degrees of Freedom	X^2	Probability of Null Hypothesis $P(H_o)$
CV	-	3	0.91	$0.99 < P_o < 0.95$
WH	I	2	3.84	$0.10 < P_o < 0.25$
	II	4	4.90	$0.25 < P_o < 0.5$
	III	4	9.83	$0.025 < P_o < 0.05^a$
LB	I	3	9.22	$0.025 < P_o < 0.05^a$
	II	3	2.62	$0.25 < P_o < 0.5$
DT	-	3	1.47	$0.50 < P_o < 0.75$
BL	I	2	0.92	$0.95 < P_o < 0.975$
	II	2	1.05	$0.50 < P_o < 0.75$
SR	-	2	0.79	$0.95 < P_o < 0.975$

a = significant at 5% level.

between two plants.

The X^2 contingency table analysis only determines that significant differences among the fractions of locations exist. It does not identify where they occur. Significant differences between two sections can be identified using the Z statistic. The Z statistic is the number of standard deviations between two fractions when the normal distribution is used as an approximation of the binomial distribution. The Z values for two section comparisons for the samples collected from system LB during June 1979 are given in table 6.11. It can be seen that there are no significant differences between the North and North Central sections or between the South and South Central section. The only significant differences occur between the sections in the Brant Beach service area and sections in the Terrace Plant service area and, of course, between the two service areas.

TABLE 6.11

ANALYSIS OF SYSTEM LB DATA FOR JUNE 1979

Section	Date	Number of Locations Sampled	Fraction with Coliforms	Z Statistic			
				North	North Central	South Central	South
North	6/7	1	0				
	6/12	10	0				
	6/14	13	0.308		0.19	1.39	0.71
	6/19	10	0		1.42	2.66[a]	2.68[a]
	6/21	9	0.111		0.60	0.36	0.50
	overall	43	0.116		0.09	1.96	2.49[a]
North Central	6/7	0	-				
	6/12	7	0				
	6/14	11	0.2	0.19		1.50	0.86
	6/19	11	0.182	1.42		1.64	1.68[b]
	6/21	20	0.050	0.60		1.17	1.30
	overall	49	0.122	0.09		1.96[a]	2.52[a]
South Central	6/7	2	0				
	6/12	16	0				
	6/14	12	0.538	1.39	1.50		0.86
	6/19	14	0.500	2.66[a]	1.64		0
	6/21	18	0.167	0.36	1.17		0.16
	overall	62	0.274	1.96[a]	1.96[a]		0.65
South	6/7	0	-				
	6/12	8	0				
	6/14	21	0.429	0.71	0.86	0.86	
	6/19	16	0.500	2.68[a]	1.86[b]	0	
	6/21	16	0.188	0.50	1.30	0.16	
	overall	61	0.328	2.49[a]	2.52[a]	0.65	
Brant Beach Service Area	6/7	1	0				
	6/12	17	0				
	6/14	24	0.292				1.47
	6/19	21	0.095				3.02[a]
	6/21	29	0.069				1.41
	overall	92	0.119				7.34[a]
Terrace Plant Service Area	6/7	2	0				
	6/12	24	0				
	6/14	33	0.485	1.47			
	6/19	30	0.500	3.02[a]			
	6/21	34	0.177	1.41			
	overall	123	0.301	7.34[a]			

a = significant at 5% level, b = significant at 10% level

The significant differences are primarily the result of a large fraction of positive samples collected from the Terrace Plant service area on 6/14/79. On that day, the high lift pumps at the Terrace Plant were out of service for three hours during the period when the samples were being collected. The pressure in the Terrace Plant service area dropped and water from the Brant Beach Plant flowed into the area causing a reversal of flow in the two large mains connecting the service areas. It can be postulated that the hydraulic disturbance in the Terrace Plant service area caused a resuspension of sediments in the mains or a shearing of bacterial growth from the walls of the mains. This hydraulic disturbance then would be responsible for the large number of positive samples collected in the South and South Central Sections on 6/14/79.

The comparisons between pairs of sections for the third sampling period for system WH are presented in table 6.12. The significant differences are between North and West Central Sections and between the Southwest and West Central Sections and are primarily the result of samples collected on 5/28/81. The samples collected on 6/4/82 were collected mostly in the Southwest and North Sections as a measure of the occurrence of coliform bacteria in the sections of the system most distant from the well. Some fire hydrants were flushed on that day to increase the flow in those distant sections. Although no samples were from the West Central section on that last sampling day, the large numbers of samples collected and the relatively high fractions of the locations with coliforms in the Southwest and North Sections contributed to the significant differences found.

It is clear that there can be significant differences in the occurrence of coliforms among the isolated sections of a distribution system or between two sections. These differences can occur on a single day and then not occur on another day of the same week. The possible occurrence of such differences should be taken into consideration in the planning of a monitoring program.

6.3.2 Differences Within Sections

The value of the concept of an isolated section of a water distribution system is that it forces recognition that there may be places where coliforms may be present and not be detected unless samples are collected in the immediate vicinity. One problem with applying the concept to some water distribution systems is that there may not be clear hydraulic separations into isolated sections. If a distribution system is one large interconnected grid how can it be divided into sections for sampling? One way of approaching this problem is to look for significant differences between adjacent sampling locations or between sampling locations two blocks apart, three blocks apart, etc.

The smallest system sampled in terms of the number of locations was the MW system which had two. The two streets were Sherwood Drive and Hall Road. The two water mains were connected together at each end and the transmission line from the well fed into the Hall Road Main. Thus, the entire distribution system was a single loop.

The coliform results from the 100ml samples collected from MW are given

TABLE 6.12

ANALYSIS OF SYSTEM WH DATA FOR MAY-JUNE 1981

Section	Date	Number of Locations Sampled	Fraction with Coliforms	Z Statistic				
				East	East Central	West Central	South-west	North
East	5/22	0	-					
	5/27	8	0		0	0	0.88	0.88
	5/28	6	0.167		1.04	0.55	0.67	0
	6/2	9	0.111		1.26	0.22	0.36	1.28
	6/4	0	-					
	Overall	23	.087		0.32	1.04	1.27	1.59
East Central	5/22	0	-					
	5/27	5	0	0		0	0.7	0.7
	5/28	6	0	1.04		0	1.54	1.04
	6/2	5	0.4	1.26		1.56	1.13	0.09
	6/4	1	0					
	Overall	17	0.118	0.32		1.35	0.83	1.12
West Central	5/22	0	-					
	5/27	13	0	0	0		1.11	1.11
	5/28	12	0	0.55	0		2.12[1]	0.55
	6/2	12	0.083	0.22	1.56		0.62	1.60
	6/4	0	-					
	Overall	37	0.027	1.04	1.35		2.46[a]	2.81[a]
South-west	5/22	1	0					
	5/27	11	0.091	0.88	0.70	1.11		0
	5/28	6	0.333	0.67	1.54	2.12[a]		0.67
	6/2	12	0.167	0.36	1.13	0.62		1.05
	6/4	13	0.308					0.07
	Overall	43	0.209	1.27	0.83	2.46[a]		0.63
North	5/22	1	0					
	5/27	11	0.091	0.88	0.70	1.11	0	
	5/28	6	0.167	0	1.04	0.55	0.67	
	6/2	8	0.375	1.28	0.09	1.60	1.05	
	6/4	27	0.296				0.07	
	Overall	53	0.245	1.59	1.12	2.81[a]	0.63	

a = significant at the 5% level

in table 6.13 divided according to location and by sampling date. The Z statistic for comparison of the fractions of samples with coliforms between the two sampling locations on a day by day basis are given in the last line of the table. It is clear that there are no significant differences between the results obtained from the two locations. The tentative conclusion from this is that there was no detectable difference in the level of contamination between the two locations. It should be noted that a great many samples were collected from each location and the level of contamination was relatively low.

The sampling results from system BL divided according to the locations and the sampling periods are presented in table 6.14. Some of the sampling location numbers are missing from the table because it was not possible to collect samples from those locations. Most of the locations which could be sampled were sampled each of the sampling periods and the number of sites sampled per location varied from 1 to 8. Samples with coliforms present were found here and there but no consistent differences among the locations are evident. Adjacent locations sometimes yielded similar results and sometimes different results and the microbiological contamination of the system could not be localized. The fraction of sites with coliform present is very small for all sections and for system BL there is no good way to group contiguous locations so that a χ^2 test for significant differences among locations can be performed. From these data it is impossible to determine if there either are or are not differences among the locations within each section. It appears that the microbiological contamination is general throughout the system but at a low level during both sampling periods. It takes a large number of samples to detect the coliforms but the locations where they are found appear to be distributed without any detectable pattern.

The sampling results for system SR divided according to the 18 locations are given in table 6.15. In this system the fraction positive was 0.125, which means that only 8 sites sampled are needed to give an expection of 1 positive site and it is easier to group contiguous locations for the χ^2 this system. The locations were grouped as is shown in the table and the results of the χ^2 test are given at the bottom of the table. No significant differences among the locations were found. This analysis is equivalent to dividing the East section into four subsections and the North section into two subsections and it is not at all surprising that the results in table 6.15 are not different from the results of the analysis by section for system SR as shown in table 6.10.

The net outcome of all this is that there is no basis, in these data, for limiting the size of an isolated section. Samples collected at one location are not necessarily similar to those collected at nearby locations nor necessarily different from those collected at locations some distance away.

6.4 HYPOTHESES ABOUT CONTAMINATION

The mechanistic approach to modeling coliform dispersion in a water distribution system is an attempt to find a method of localizing contamination in the system. If this approach were to prove to be successful, it

TABLE 6.13

COMPARISON OF SAMPLING RESULTS (100ml SAMPLES) FROM THE TWO LOCATIONS OF SYSTEM MW

Location	Coliform Sampling Results on Indicated Date 10/8/80	10/23/80	10/30/80	11/6/80	11/11/80	11/13/80	11/19/80	12/11/80	Overall
Sherwood Drive									
Sites Sampled	4	2	10	10	10	13	10	10	33
Number of Samples	8	4	20	20	20	68	40	40	220
Fraction Positive	0	0.25	0.15	0.05	0	0.19	0	0.02	0.086
Mean Density (per 100ml)	0	0.25	0.25	0.40	0	1.56	0	0.02	0.55
Hall Road									
Sites Sampled	0	2	3	9	6	1	4	6	18
Number of Samples		4	6	18	12	4	16	24	84
Fraction Positive		0.25	0	0	0.08	0.25	0	0.04	0.048
Mean Density (per 100ml)		0.25	0	0	0.08	0.25	0	0.12	0.071
Z Values for Comparisons of Fractions Positive		0	1.61	0.96	1.31	0.29	0	0.19	1.12

TABLE 6.14

COMPARISON OF SAMPLING RESULTS (100ml SAMPLES) FROM 68 SAMPLING LOCATIONS OF SYSTEM BL

Location Number	March-April 1980			June 1981		
	Sites Sampled	Fraction Positive	Mean Density (per 100ml)	Sites Sampled	Fraction Positive	Mean Density (per 100ml)
102	8	0.125	0.06	1	0	-
103	3	0	-	1	0	-
104	3	0	-	4	0	-
105	6	0	-	2	0	-
106	8	0	-	3	0.333	0.67
107	3	0	-	4	0	-
108	3	0	-	4	0	-
108	3	0	-	4	0	-
109	3	0.333	>26.67	1	0	-
111	3	0.333	0.17	2	0	-
112	3	0	-	2	0	-
113	2	0	-	3	0	-
114	4	0	-	3	0	-
115	0	-	-	1	0	-
116	2	0	-	1	0	-
117	2	0	-	1	0	-
118	2	0	-	2	0.500	0.75
119	1	0	-	2	0	-
120	4	0	-	2	0	-
121	3	0	-	0	-	-
122	4	0	-	3	0	-
123	3	0	-	3	0.333	0.67
124	3	0	-	0	-	-
125	3	0	-	1	0	-
126	1	0	-	0	-	-
127	3	0.333	0.17	3	0.333	0.33
128	1	0	-	1	0	-
Total for Southeast Section	81	0.049	>1.01	50	0.080	0.11
132	2	0	-	3	0.333	0.17
133	5	0	-	2	0	-
134	1	0	-	3	0	-
135	2	0	-	0	-	-
136	2	0	-	1	0	-
137	2	0	-	0	-	-
138	2	0	-	1	0	-
139	3	0	-	1	0	-
140	4	0	-	1	0	-
141	3	0	-	3	0	-

TABLE 6.14 (continued)

	March-April 1980			June 1981		
Location Number	Sites Sampled	Fraction Positive	Mean Density (per 100ml)	Sites Sampled	Fraction Positive	Mean Density (per 100ml)
142	4	0	-	0	-	-
143	3	0	-	2	0	-
144	3	0	-	4	0	-
145	4	0	-	5	0.200	0.60
146	3	0.333	0.17	2	0.500	0.25
147	3	0	-	3	0.333	0.33
148	5	0	-	5	0	-
149	3	0	-	0	-	-
150	6	0	-	2	0	-
151	4	0	-	2	0	-
151	4	0	-	2	0	-
152	2	0	-	5	0.200	0.40
153	2	0.500	0.25	1	0	-
154	5	0	-	5	0.400	0.40
155	4	0	-	2	0	-
156	7	0	-	3	0.333	0.33
Total for Central Section	84	0.024	0.01	56	0.142	0.18
158	1	0	-	4	0	-
159	1	0	-	3	0	-
160	5	0	-	4	0.500	15.62
161	3	0	-	4	0	-
162	3	0	-	4	0	-
163	3	0	-	3	0.333	0.17
164	5	0	-	3	0.333	0.17
165	3	0	-	7	0	-
167	7	0	-	1	0	-
168	4	0.250	0.12	2	0.500	0.25
169	4	0	-	6	0	-
170	8	0	-	7	0.142	0.07
171	6	0	-	6	0	-
172	6	0.167	0.08	5	0	-
173	5	0	-	3	0.333	1.00
174	6	0	-	2	0	-
Total for Northeast Section	70	0.029	0.01	64	0.109	1.06

TABLE 6.15

COMPARISON OF SAMPLING RESULTS (100ml SAMPLES) FROM 18 SAMPLING LOCATIONS OF SYSTEM SR

Location Number	Sites Sampled	Fraction Positive	No. of Positive Sites (Grouped Locations)	
			Expected	Observed
101	4	0	1.125	0
102	5	0		
103	5	0	1	0
104	3	0		
105	9	0.111	1.125	1
106	0			
107	7	0.143	2.75	3
108	15	0.133		
109	24	0.167	3.875	4
110	1	0		
111	6	0		
Total for East Section	79	0.101		
112	4	0.500	5.5	6
113	40	0.100		
Total for West Section	44	0.136		
114	3	0	1.12	0
115	3	0		
116	3	0		
117	22	0.181	2.75	4
118	22	0.181	2.75	4
Total for North Section	53	0.151		
χ^2 =5.11 for 8 df			$.50<P(H_o)<.75$	

would be possible to work out a sampling plan which would be superior to random selection of sampling sites. That is, a sampling site could be selected which would represent a main or a part of an isolated section and other samples from that main or part would not be required. Achieving such a goal would greatly increase the efficiency of sampling.

The hypotheses discussed in this section could be used to formulate sampling plans which would be very efficient for detecting contamination and for localizing such contamination. Unfortunately, none of the hypotheses were verified by the data collected during this project and thus it is not possible to construct a sampling model using the mechanistic approach. Since all of the systems sampled during this project were small, it is possible that some of the hypotheses might be true for larger systems.

All of the hypotheses were tested using data from all nine of the systems sampled. However, since none of these hypotheses were found to be valid for any of the systems, it seems unnecessary to present all of the analyses which were made. Data from the sampling of system BL during June 1981 are used as examples of the results obtained. Some data from other systems are also used to illustrate particular points.

In order to test these hypotheses, it was necessary to construct flow diagrams for the sampling locations of each system. The flow diagram for the isolated sections of system BL is presented in figure 6.3. In constructing these diagrams it was sometimes necessary to make assumptions about the direction of flow in a particular street lateral. The assumptions were based on the concepts that flow is normally from the street laterals closer to the treatment plant to the ones further away and from the larger pipes to the smaller pipes. Using these concepts, the shortest distance of flow from the treatment plant to any street lateral was selected.

6.4.1 Upstream and Downstream Samples

The first of these hypotheses tested was that samples collected downstream from a site from which a positive sample had been obtained would also have coliforms present. The companion hypothesis is that samples collected upstream from a site where a negative sample would also be negative. These hypotheses should be tested using data collected on a single day because sources of contamination may not presist for more than a short time. It is clear that if one of these hypotheses were true, the other one would also be true.

The locations of the samples with coliforms collected from system BL in June 1981 are listed in table 6.16 along with all downstream locations sampled. The downstream locations are divided into those where coliforms were found and those where coliforms were not found. From this table it is clear that negative samples are frequently collected downstream from positive samples and, conversely, that positive samples are often collected upstream from negative samples. The same conclusion was reached by examination of similar data from other systems. Thus, the two hypotheses have to be rejected, at least in the general case.

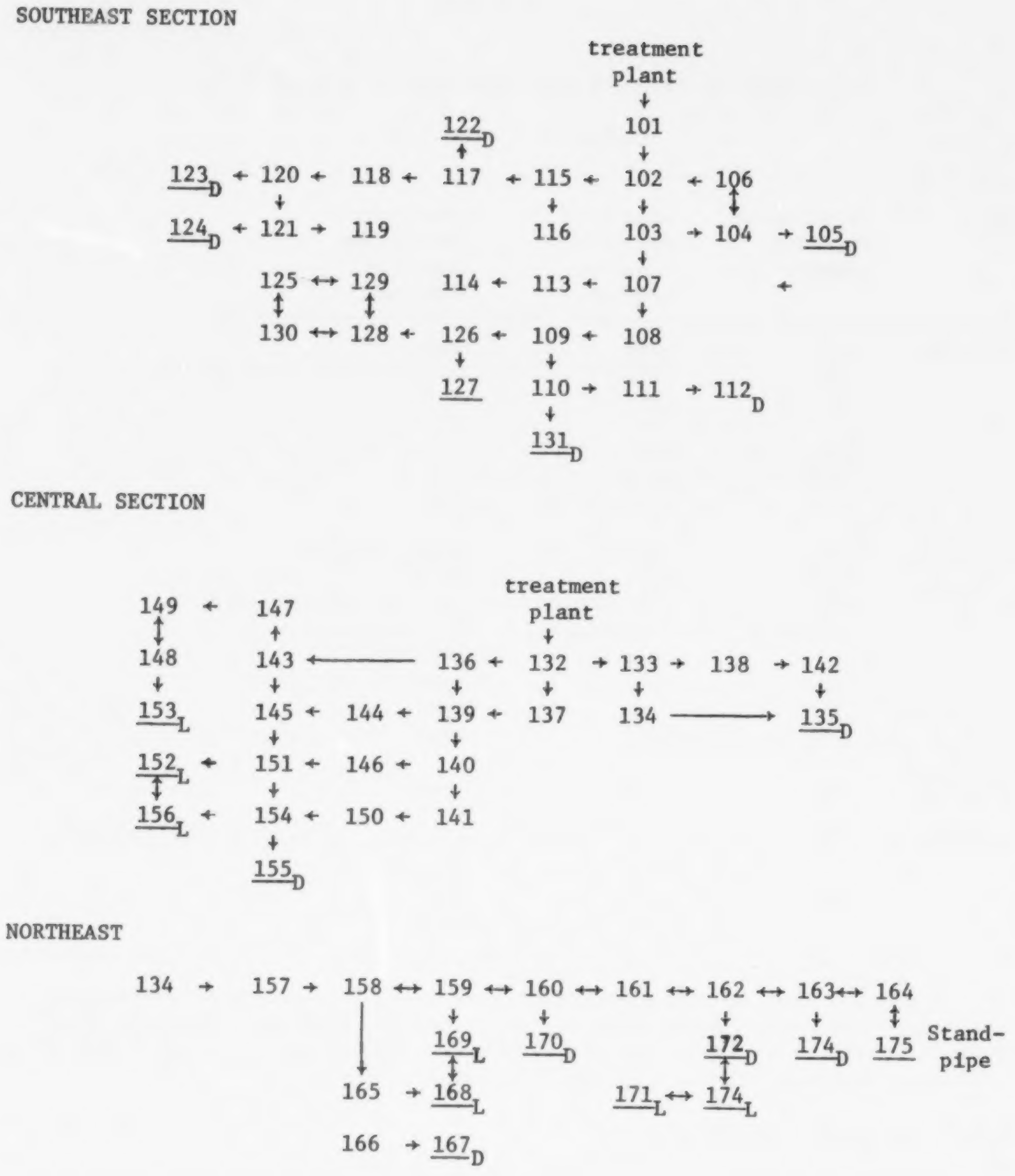

SOUTHEAST SECTION
treatment plant
101
122D
123D 120 118 117 115 102 106
124D 121 119 116 103 104 105D
125 129 114 113 107
130 128 126 109 108
127 110 111 112D
131D
CENTRAL SECTION
treatment plant
149 147
148 143 136 132 133 138 142
153L 145 144 139 137 134 135D
152L 151 146 140
156L 154 150 141
155D
NORTHEAST
134 157 158 159 160 161 162 163 164
169L 170D 172D 174D 175 Stand-pipe
165 168L 171L 174L
166 167D

Figure 6.3 Flow Diagrams for the Isolated Sections of System BL.

TABLE 6.16

LOCATIONS OF POSITIVE COLIFORM SAMPLES COLLECTED FROM SYSTEM BL - JUNE 1981

Date	Location of Positive Sample	Locations of Downstream Positive Samples	Locations of Downstream Negative Samples
6/16/81	122	none	none
	160	163,164	161,162,170A,170B,171,172,174
	163	164	
	164	none	
	152B	none	156
	154	none	156
6/18/81	118	none	120A,120B,123
	127C	none	none
	145	154,156A	148,152,153,155,156B
	146	154,156A	155,156B
	154	156A	155,156B
	156A	none	156B
	168	none	169A,169B
	170	none	none
6/23/81	106	none	none
6/25/81	132	147	133,134,136,143,144,145,148,151, 152,155
	147	none	148
	173	none	none

A modification of these two hypotheses would be that positive samples occur more frequently downstream of other positive samples. However, there were no instances in which enough samples were collected on a single day to provide a reasonable basis for testing that hypothesis.

6.4.2 Peripheral Locations

Another hypothesis tested is that coliform bacteria are found more frequently in samples from peripheral locations than in samples from other locations. Table 6.17 gives comparisons of the fractions of peripheral and other (nonperipheral) locations positive for coliforms for the various systems and sampling periods. System MI was omitted because there were too few positive samples for a reasonable test and system MW was omitted because the entire system consists of two locations connected to for a loop. The Z statistic is based on the use of the normal distribution to approximate the binomial distribution and is a measure of the number of standard deviations between the two fractions positive. A Z value of 1.96 would indicate a difference significant at the 5% level. No significant differences were found. Actually, the fraction positive for the peripheral locations was greater than the fraction positive for the other locations in only 7 of the 11 cases.

When the comparisons were made for the isolated sections and for individual sampling days, there were a few significant differences between peripheral and other locations. For instance, in system SR all the positive samples collected in the North Section were from peripheral locations and all the positive samples collected in the East Section were from non-peripheral locations. However, when the comparisons were made for the entire system the results from the individual sections balanced out. There is clearly no basis for concluding that peripheral locations are more likely to yield coliforms when the systems are considered as a whole.

TABLE 6.17

PERIPHERAL VERSUS OTHER LOCATIONS

System-Sampling Period	Peripheral Locations		Other Locations		Z Statistic
	No. of Locations Sampled	Fraction Positive	No. of Locations Sampled	Fraction Positive	
CV	70	0.029	145	0.055	0.85
WH-I	28	0.107	38	0.079	0.39
WH-II	54	0.463	100	0.360	0.76
WH-III	55	0.127	118	0.169	0.65
LB-I	87	0.195	128	0.250	1.38
LB-II	92	0.369	140	0.407	0.58
DT	57	0.053	117	0.051	0.06
BL-I	73	0.055	162	0.025	1.17
BL-II	54	0.167	116	0.086	0.86
SR	71	0.155	105	0.105	1.05
BG	24	0.250	29	0.172	0.70

6.4.3 Distance from Treatment Plant

Another hypothesis of the proposed mechanistic approach is that samples with coliform are more likely to be found far from the treatment plant (source of water to the distribution system) than close to the treatment plant. The concept behind this hypothesis is that the microbiological contamination is introduced into the water after it leaves the plant and the further the water has traveled, the greater the chance it has become contaminated.

TABLE 6.18

ANALYSIS OF COLIFORM OCCURRENCE AS A FUNCTION OF DISTANCE FROM THE TREATMENT PLANT SYSTEM BL - JUNE 1981 DATA

Distance from Plant (blocks)	Potential Sampling Locations	No. of Sites Sampled	Fraction Positive
1	2	3	0.333
2	4	3	0
3	8	12	0.083
4	11	27	0.074
5	13	37	0.081
6	11	26	0.115
7	11	29	0.172
8	8	16	0
9	4	11	0.091
10	2	6	0.333

The coliform data from the 1981 sampling of system BL are presented in table 6.18. It is clear that the fraction of the sites which yielded samples with coliforms does not increase with distance from the treatment plant. Indeed, in some instances, locations nearer the treatment plant yielded more positive samples and more coliform bacteria than more distant locations. However, system BL has an elevated storage tank which is eleven blocks from the treatment plant which could be considered as a "source" of water to the distribution system and which could affect the microbiological water quality in the northeast and southeast sections of the distribution system. Thus, system BL may not provide a good test of this hypothesis.

Systems WH and LB both have finished water storage in standpipes close to the treatment plants. Thus, they should provide adequate tests of the hypothesis. Data from these two systems are presented in tables 6.19 and 6.20. Again it is clear that the frequency of occurrence of samples with coliforms does not increase with distance from the water source. The same type of analysis was made of the occurance of coliforms in the other systems and no evidence supporting the hypothesis was found.

TABLE 6.19

ANALYSIS OF COLIFORM OCCURRENCE AS A FUNCTION OF DISTANCE FROM THE TREATMENT PLANT
SYSTEM WH MAY-JUNE 1981

Distance from Plant (blocks)	Potential Sampling Locations	No. of Sites Sampled	Fraction Positive
1	8	7	0
2	11	16	0.125
3	22	25	0.040
4	11	18	0.111
5	11	11	0.182
6	2	2	0.500
7	5	8	0
8	7	10	0.200
9	6	12	0.167
10	6	10	0.300
11	8	16	0.375
12	8	18	0.167
13	7	10	0.200
14	2	4	0.500
15	2	6	0.167

TABLE 6.20

ANALYSIS OF COLIFORM OCCURRENCE AS A FUNCTION OF DISTANCE FROM THE TREATMENT PLANT
SYSTEM LB

Distance from Plant (blocks)	Potential Sampling Locations	No. of Sites Sampled	Fraction Positive
0 - 5	120	98	0.357
6 - 10	99	66	0.378
11 - 15	88	43	0.326
16 to 20	108	58	0.345
21 - 25	106	67	0.373
26 - 30	75	79	0.304
>30	48	36	0.278

6.5 SUMMARY AND IMPLICATIONS

The concept of dividing a water distribution system into hydraulically isolated sections for monitoring purposes is important. Coliform bacteria in a water distribution system travel with the flow of water not against it. Thus, if they are introduced into a system at only one point, they would be found downstream of that point and not upstream of it. Clearly, there are places in a water distribution system where coliforms can be present and they will not be detected by samples taken from other parts of the system.

The concept of classifying isolated sections of a water distribution system according to their geometry is also important. Flow reversals are possible and probably occur in the pipes which make up grid or loop sections depending upon the location of the water demand; thus, it is not possible to determine upstream and downstream directions in such sections in any absolute sense. On the other hand, flow reversals cannot occur in the pipes which make up linear or dendritic sections unless there is more than one source of flow into the section. Such factors as direction of flow and reversal of flow should be considered in the selection of "representative" sampling locations.

In testing for differences among sections, the small fraction of sites with coliforms present coupled with a relatively low number of samples spread over several sections was a problem for the analysis of data from systems CV and DT and from the first sampling period for system WH. This problem was circumvented by grouping together data from contiguous sections. This grouping of data was necessary but it somewhat weakens the purpose of the test.

Significant differences among the sections were found for the data sets for the third sampling period for system WH and for the first sampling period for system CV. This shows that such differences can exist at times and thus that monitoring results may be biased if all sections of a distribution system are not included in the monitoring program.

No significant differences were found among the locations within the sections of a system in the fraction of the sites with coliforms present. Actually, the χ^2 contingency table test could not be used for testing for differences among locations for most systems because there were few samples per location and a relatively low fraction of samples positive for coliforms. The contingency table test was used for system SR by grouping results of some locations. The conclusion from these results is that there is no basis in any of the data collected from these systems for subdividing hydraulically isolated sections.

If coliform contamination in a water distribution system was localized in one part of the system or in one particular type of location, the mechanistic approach could be of value in determining where to sample to find the coliforms. In this section, we have presented the results of testing the hypotheses that if coliforms are found in a particular location they should be found in all downstream locations, that coliforms are more likely to occur in peripheral than in non-peripheral locations, and that the frequency of occurrence of coliforms increases with distance of flow from the water source into the system. No evidence to support any of these hypotheses was found. The conclusion from this is that in planning a monitoring program all parts of the system and all types of locations should be included. The best method of selecting locations for sampling is a randomization procedure which gives each potential sampling location the same probability of being collected.

In the water works industry it is commonly assumed that if coliforms are not found in the finished water leaving the treatment plant and are not found in locations the most distance from the treatment plant, they will not occur at intermediate locations. Our results demstrate that this is not true, at least, for the systems we sampled. The implications of this finding is that all parts of the distribution system have to be included in the monitoring program in order to have the samples collected at points which are representative of the conditions with the distribution system.

REFERENCES

Aitchison, J. and J. A. C. Brown. 1957. *The Lognormal Distribution*. Cambridge University Press, Cambridge.

APHA. 1976. *Standard Methods for the Examination of Water and Wastewater* 14th Ed., American Public Health Association, Washington, D.C.

Bordner, R., J. Winter, and P. Scarpino. 1978. *Microbiological Methods for Monitoring the Environment* - *Water and Wastes*, EPA-600/8-78-017, Environmental Monitoring and Support Laboratory, U.S. Environmental Protection Agency, Cincinnati, Ohio.

Brodsky, M. H., B. W. Ciebin, and D. A. Schiemann. 1978. Simple bacterial preservation medium and its application to proficiency testing in water bacteriology. *Appl. Environ. Microbiol.*, *35*:487-491.

El-Shaarawi, A. H., S. R. Esterby, and B. J. Dutka. 1981. Bacterial density in water determined by Poisson or negative binomial distributions. *Appl. Environ. Microbiol.*, *41*:107-116.

Ewing, W. H. and M. A. Fife. 1972. *Enterobacter* *agglomerans* (Beijerinck) Comb. Nov. (the Herbicola-Lathyri Bacteria). *Int. Jour. Systm. Bact.*, 22:4-11.

Geldreich, E. E.,1971. Application of bacterial data in potable water surveillance. *Jour. Am. Water Works Assoc.* *63*:4:225-229.

Geldreich, E. E., 1975. *Handbook for Evaluating Water Bacteriological Laboratories*, 2nd Ed., EPA-670/9-75-006, Municipal Environmental Research Laboratory, U.S. Environmental Protection Agency, Cincinnati, Ohio.

Geldreich, E. E., M. J. Allen, and R. H. Taylor. 1978. Interferences to coliform detection. pp. 13-20. *In*. *Evaluation of the Microbiology Standards for Drinking Water* (C. W. Hendricks, ed.), U. S. Environmental Protection Agency, Washington, D. C.

Gerba, Charles, P., C. Wallis, and J. L. Melnick. 1975. Microbiological hazards of household toilets: droplet production and the fate of residual organisms. *Appl. Microbiol.* *30*:2:229-237.

Goshko, M. A., H. A. Minnigh, W. O. Pipes, and R. R. Christian. 1981. Relationships between standard plate counts and other parameters in water distribution systems. Presented at the 1981 Annual Conference. American Water Works Association, St. Louis, Mo., June 1981.

Harrison, J. 1978. Experiences with the coliform standard. pp. 159-168. In. Evaluation of the Microbiology Standards for Drinking Water (C. W. Hendricks, ed.), U. S. Environmental Protection Agency, Washington, D. C.

Lariviere, J. W. 1978. Survival of Escherichia coli in four diluents. M.S. thesis, Department of Biological Sciences, Drexel University, Philadelphia, Pa.

LeChevalier, M. W., R. J. Seidler, and T. M. Evans. 1980. Enumeration and characterization of standard plate count bacteria in chlorinated and raw water supplies. Appl. Environ. Microbiol. 40:922-929.

Muenz, L. R. 1978. Some statistical considerations in water quality control. pp. 49-56. In. Evaluation of the Microbiology Standards for Drinking Water (C. W. Hendricks, ed.), U. S. Environmental Protection Agency, Washington, D. C.

Pielou, E. C. 1969. An Introduction to Mathematical Ecology, Wiley-Interscience, New York.

Pipes, W. O., P. Ward, and S. H. Ahn. 1977. Frequency distributions for coliform bacteria in water. Jour. Am. Water Works Assoc., 69:644-647.

Pipes, W. O. and R. R. Christian. 1978. A sampling model for coliforms in water distribution systems. Presented at the 1978 Annual Conference, American Water Works Association, Atlantic City, N.J., June 1978.

Sokal, R. R., and F. J. Rohlf. 1969. Biometry. W. H. Freeman and Co., San Francisco, Ca.

Thomas, H. A., Jr. 1952. On averaging results of coliform tests, Jour. Boston Soc. Civil Engineers, 39:253-264.

Thomas, H. A., Jr. 1955. Statistical analysis of coliform data. Sew. Ind. Wastes, 27:212-228.

U. S. E. P. A. 1974. Methods for Chemical Analysis of Water and Wastes. Office of Water Supply, U. S. Environmental Protection Agency, EPA-625/6-74-003. Washington, D. C.

U. S. E. P. A. 1976. National Interim Primary Drinking Water Regulations. Office of Water Supply, U. S. Environmental Protection Agency, EPA-570/9-76-003. Washington, D. C.

U. S. Public Health Service. 1962. Drinking Water Standards. Department of Health, Education, and Welfare., Washington, D. C.

Velz, C. J. 1951. Graphical approach to statistics. IV. Evaluation of bacterial density. Water Sew. Works, Feb. 1951, pp. 66-73.

APPENDIX J

EFFECT OF FILTER BACKWASHING ON BACTERIOLOGICAL QUALITY OF DRINKING WATER

Report of Special Study

APPENDIX J

EFFECT OF FILTER BACKWASHING ON BACTERIOLOGICAL QUALITY OF DRINKING WATER

J.1 OBJECTIVE

To determine if there are differences in the microbial quality of water as measured by coliform and standard plate counts entering a distribution system between when the filters are being backwashed and when they are operating normally.

J.2 HYPOTHESES

The purpose of water filtration is to remove particulate matter, including bacteria, from the water before it enters a distribution system. These solids are separated from the filter by backwashing but some small amounts may be left in the filter after backwashing is complete and the filter medium is allowed to settle back into place. During backwashing the total amount of water being produced by the treatment plant is reduced and finished water from storage may enter the distribution system.

1. If prechlorination is not used or is not completely effective, the filter backwash water may contain large numbers of bacteria, some of which will be coliforms.

2. After backwashing and after the filter medium has been allowed to settle back into place, the first flush of water through the filter may contain large numbers of bacteria including some coliforms.

3. Some coliforms may enter water distribution systems from filtered water reservoirs, standpipes or elevated storage tanks if the pressure in the distribution system is allowed to drop during filter backwashing.

J.3 SYSTEMS STUDIED

The water systems selected for study were Downingtown, Pennsylvania (sampled January and February 1980) and Brooklawn, New Jersey (sampled March and April 1980). In addition some samples were collected in July 1979 during filter backwashing at the two treatment plants used by the Long Beach Water Company. Although these data were collected before the backwash study was designed they are included as additional evidence related to two of the hypotheses being tested. Basic descriptive information on the treatment plants and their water sources is given in Appendicies D, E, and C.

J.3.1 Long Beach Water Company

The two treatment plants are 3 miles apart and serve different areas of the distribution system. However, they are essentially identical facilities. During the summer, when these samples were collected, the filters are backwashed on Mondays, Wednesdays, and Fridays, those at the Terrace Plant in the morning and those at the Brant Beach Plant in the afternoon. The backwash water is discharged into the sewer system. There are three high lift pumps at each plant rated at 1000 gpm, 500 gpm, and 500 gpm. During normal operation the 1000 gpm pump is used to pump water through the filters into the distribution and is turned on and off by a pressure switch located at the base of the elevated storage reservoir or standpipe. During the winter the high lift pump operates about 2.5 hours per day. During backwashing the 1000 gpm pump is used to backwash one filter while the two 500 gpm pumps are used to pump water through the other two filters into the distribution system. Since the backwash water is taken from the distribution system, the water use for backwashing must come from the elevated storage. There is no level or pressure recorder on either the standpipe at the Terrace Plant or the elevated storage tank at the Brant Beach Plant and the water pumpage recorder is turned off during backwashing. Thus, there is no documentation of water use during backwash nor of the drawdown of water in storage.

J.3.2 Downingtown

The water filtration plant at Downingtown is operated only 16 hours per day. The filtered water in the 3.75 million gallon open reservoir is pumped into the distribution system by high lift pumps which are operated independently of the filtration plant. Chlorine to provide a disinfectant residual in the treated water is added at the high lift pumping station. The elevated storage tanks are on the distribution system at maximum distances from the treatment plant and contain enough water to supply the system for 2 to 4 days even if the high lift pump were not operating. There are no loss-of-head meters on the filters. The filters are backwashed after 30 hours of operation using water from an elevated backwash water tank. The backwash water is discharged to a settling pond from which the overflow goes to the raw water storage.

J.3.3 Brooklawn

The operation of the Brooklawn water treatment plant is similar to that of the two plants of the Long Beach Water Co. The high lift pump supplies pressure to force the water through the filters and into the distribution system. It is turned on and off by a pressure switch on the main which supplies water to the distribution system. During backwashing, the high lift pump is turned off and water from the distribution system is used for backwash. The three filters are backwashed every morning for 10 minutes each and the drawdown in the elevated storage tank on the other side of town is only 1 to 2 feet. The backwash water is discharged into a storm sewer which empties into a tidal flat area of the Delaware estuary.

J.4 RESULTS OF TESTING HYPOTHESES

J.4.1 Bacteria in Backwash Water

J.4.1.1 Long Beach Water Company

Backwash water samples were not collected at either plant because this sampling was carried out before the backwash study was designed. However, there is no disinfection of the water before filtration so the hypothesis could not be tested properly by samples from these plants. Bacteriological results for raw water samples are given in Table J1.

The one confirmed coliform found in the raw water at the Brant Beach Plant on 6/14 was identified as *Citrobacter freundii*. A presumptive coliform was found in one of the samples collected on 6/19 but it did not confirm and was identified as *Enterobacter agglomerans*. The standard plate counts were consistently low except for one sample collected on 6/21.

No presumptive coliforms were found in the raw water at the Terrace Plant. The standard plate counts were more variable than those on samples from the Brant Beach Plant. However, it is clear that the well water is not a major source of bacteria.

J.4.1.2 Downingtown Water Treatment Plant

The results pertinent to the first hypothesis are presented in Table J2. The volumes of raw water filtered for the coliform determinations were 1 ml and 10 ml. Even so, some filters on samples from 1/15 had more than 80 coliform colonies present. Coliforms identified from the raw water included *Escherichia coli*, *Citrobacter freundii*, *Klebsiella oxytoca*, and *Enterobacter* spp.. The standard plate counts on the raw water were also very high.

It is clear that chlorination, flocculation and settling at this plant is very effective in reducing the coliform density and standard plate count. The one coliform found in the treated water samples was *Klebsiella pneumoniae*.

The multiple tube fermentation tests on the backwash water did not provide definite evidence of the presence of coliforms. The tests were run using standard 10 ml portions on the first three sampling days and no positive tubes were found. On the last two sampling days five 100 ml, five 10 ml, and five 1 ml portions were used. Gas production was found in some of the 100 ml portions but it was not possible to confirm the tests or isolate coliforms from the broth. One isolate from one backwash sample collected on 1/29 was identified as *Enterobacter agglomerans* but it did not produce gas in brilliant green bile broth. The standard plate counts on the backwash water were not significantly different from those on the treated water samples.

J.4.1.3 Brooklawn Water Treatment Plant

These results are presented in Table J3. No coliforms were found in the

TABLE J1

RESULTS OF BACTERIOLOGICAL TESTS ON RAW WATER SAMPLES
LONG BEACH WATER COMPANY

Date	Brant Beach Plant					Terrace Plant				
	No. of Samples	Coliforms (per 100ml)		Std.Plate Count (per ml)		No. of Samples	Coliforms (per 100ml)		Std. Plate Count (per ml)	
		Mean	Std. Dev.	Mean	Std. Dev.		Mean	Std. Dev.	Mean	Std. Dev.
6/7/79	2	0	-	0.25	0.35	2	0	-	0.25	0.35
6/12/79	2	0	-	0.5	0.5	2	0	-	15	21.2
6/14/79	2	0.5	0.5	0.25	0.35	2	0	-	0.5	0
6/19/79	26	0	-	1.69	6.26	20	0	-	0.1	0.26
6/21/79	2	0	-	7.25	10.25	2	0	-	0	-
7/10/79	2	0	-	0	-	2	0	-	0	-
7/12/79	2	0	-	0.5	0.5	2	0	-	30.5	23.3
7/17/79	2	0	-	0	-	2	0	-	0	-
7/18/79	1	0	-	1	-	3	0	-	13.3	23.1
7/19/79	2	0	-	0	-	2	0	-	0	-
Overall	43	0.02	0.02	1.45	5.29	39	0	-	3.45	10.81

TABLE J2

RESULTS OF BACTERIOLOGICAL TESTS ON WATER SAMPLES BEFORE FILTRATION
DOWNINGTOWN, PENNSYLVANIA

Date	Raw Water					Settled Water					Backwash Water			
	No.of Samp.	Coliforms (per 100ml)		Std.Plate Count(per ml)		No.of Samp.	Coliforms (per 100ml)		Std.Plate Count(per ml)		No.of Samp.	Presumptive MPN (per 100ml)	Std.Plate Count(per ml)	
		Mean	Std. Dev,	Mean	Std. Dev.		Mean	Std. Dev.	Mean	Std. Dev.			Mean	Std. Dev.
1/8/80	6	298	281	27.3	21.6	8	0	0	2.31	1.58	5	< 2	1.0	1.24
1/15/80	6	TNTC	-	TNTC	-	6	1	0.17	31.0	15.9	5	< 2	29.0	27.22
1/22/80	6	858	294	TNTC	-	6	0	0	21.8	26.1	5	< 2	15.6	7.37
1/29/80	6	1450	459	600	161	5	0	0	22.2	13.4	5	2.3	28.2	14.65
2/5/80	6	533	197	TNTC	-	6	0	0	54.8	70.3	5	2.3	9.9	8.00
Overall	30	TNTC	-	TNTC	-	31	0.03	0.03	24.5	34.7	25		12.83	13.96

TABLE J3

RESULTS OF BACTERIOLOGICAL TESTS ON WATER BEFORE FILTRATION
BROOKLAWN, NEW JERSEY

Date	Raw Water					Treated Water					Backwash Water				
	No.of Samp.	Coliforms (per 100ml)		Std.Plate Count(per ml)		No.of Samp.	Coliforms (per 100ml)		Std.Plate Count(per ml)		No.of Samp.	Presumptive MPN (per 100ml)		Std.Plate Count (per ml)	
		Mean	Std. Dev.	Mean	Std. Dev.		Mean	Std. Dev.	Mean	Std. Dev.		Geometric Mean	Std.Dev.	Mean	Std. Dev.
3/4/80	6	0	-	4.92	4.29	10	0	-	0.25	0.264	5	0.83	2.47	3.2	2.39
3/11/80	10	0	-	2.80	4.34	10	0	-	0.40	0.966	5	1.22	1.78	3.4	4.88
3/18/80	10	0	-	2.10	3.14	10	0	-	0.20	0.422	5	2.3	0	0.2	0.45
3/25/80	10	0	-	1.0	2.0	10	0	-	0.20	0.422	5	<0.2	-	0	-
Overall	36	0	-	2.43	3.59	40	0	-	0.26	0.566	20	0.83	2.84	1.7	2.99

well water or treated water samples. The standard plate count for the treated water is consistently lower than that of the raw water. The 10 mg/l chlorine dose added after aeration and before filtration disappears rapidly in the filters. Free chlorine levels in the treated water and filtered water samples on 3/4, 3/11, and 3/18 ranged between 0.4 and 0.7 mg/l. On 3/25 the free chlorine concentration in the filtered water was 1 mg/l. It appears that the variation in the chlorine residual is related to variable chlorine demand in the well water and a reduction in the standard plate count is achieved in spite of the chlorine demand.

The multiple tube fermentation test of the backwash water samples consisted of five 100 ml tubes, five 10 ml tubes, and five 1 ml tubes. No gas production was detected in any of the 1 or 10 ml tubes. Gas production was detected in all of the 100 ml tubes for all five samples on 3/18 but in none of the 100 ml tubes on 3/25. Variable numbers of 100 ml tubes showed gas production in the different samples on 3/4 and 3/11. Transfers from the LTB tubes with gas to BGLB did not result verification of the presence of coliforms. An isolate from one of the 100ml tubes from 3/19 was identified as *Klebsiella ozaenae* and required 5 days to produce gas in BGLB. No other isolates were members of the Enterobacteriacae.

J.4.2 Coliforms in Filtered Water

None of the three systems tested filter to waste following backwash. The treatment plants using well water take backwash water out of the distribution systems and filtering to waste would allow further drawdown of elevated water storage. Since the filtered water at the Downingtown Water Treatment Plant passes through an open reservoir and is chlorinated before it is pumped into the distribution system, there is no particular reason to filter to waste there.

J.4.2.1 Long Beach Water Company Plants

Samples of the water entering the distribution system during backwash were collected on 7/18/79. Samples of water entering the distribution system while the filters were being operated normally were collected on other sampling days. A comparison of the bacteriological results of these two different sets of samples is given in Table J4. No coliforms were found in the finished water samples and the standard plate counts are lower during backwashing than at other times.

J.4.2.2 Downingtown Water Treatment Plant

The bacteriological results on filtered water samples are given in Table J5. The volumes filtered were 1000ml each on 1/8 and 500 ml each on the other sampling days. The turbidity of all of these samples was very low, < 0.1 NTU, but the time required for filtration of the 1000 ml samples was 10 to 15 minutes each. The 500 ml samples filtered in about 2 minutes each.

Two presumptive coliforms were found in the filtered water samples after backwashing on 1/8. One did not produce gas in BGLB and did not grow on EMB

TABLE J4

RESULTS OF BACTERIOLOGICAL TESTS ON FILTERED WATER SAMPLES
LONG BEACH WATER COMPANY

Plant	During Backwashing					Other Samples				
	No.of Samples	Coliforms (per 100ml)		Std.Plate Count (per ml)		No.of Samples	Coliforms (per 100ml)		Std.Plate Count (per ml)	
		Mean	Std. Dev.	Mean	Std. Dev.		Mean	Std. Dev.	Mean	Std. Dev.
Brant Beach	6	0	-	0.17	0.41	29	0	-	1.76	3.93
Terrace	17	0	-	2.35	2.35	26	0	-	3.48	11.37

TABLE J5

RESULTS OF BACTERIOLOGICAL TESTS ON FILTERED WATER SAMPLES
DOWNINGTOWN, PENNSYLVANIA

Date	Before Backwash					After Backwash				
	No.of Samp.	Coliforms (per 100ml)		Std. Plate Count (per 100ml)		No.of Samp.	Coliforms (per 100ml)		Std.Plate Count (per 100ml)	
		Mean	Std. Dev.	Mean	Std.Dev.		Mean	Std. Dev.	Mean	Std.Dev.
1/8/80	3	0	-	0.667	0.577	10	0.01	0.032	0.05	0.158
1/15/80	15	0	-	8.33	5.55	15	0	-	6.30	2.21
1/22/80	15	0	-	4.92	2.59	15	0	-	4.23	1.23
1/29/80	13	0	-	2.54	1.97	12	0	-	0.88	1.11
2/5/80	15	0	-	4.79	1.60	15	0	-	4.64	2.89
Overall	61	0	-	4.70	4.00	67	1	0.015	3.44	2.99

agar and is not recorded as a verified coliform. The other did verify in BGLB and was identified by the API system as a strain of Enterobacter which was never found in the raw water. The most likely source for these bacteria is the backwash water storage tank.

No significant differences in the standard plate count in the filtered water before and after filtration were found.

J.4.2.3 Brooklawn Water Treatment Plant

The bacteriological results on filtered water samples are presented in Table J6. The volumes filtered were 500 ml. No coliforms were detected in the filtered water before or after filtration which is consistent with the results on the raw water and treated water samples. The standard plate counts on samples before and after filtration are not significantly different from each other nor from the standard plate counts on the treated water before filtration (Table J3).

J.4.3 Coliforms from Elevated Storage Tanks

J.4.3.1 Long Beach Water Company

There is no record of how much drawdown of water in elevated storage occurs during backwashing. The water pumpage at each of the two plants on week days during July 1979 varied between 900,000 and 1,150,000 gallones. If it is assumed that the water useage is distributed evenly throughout the day and the water going into the distribution system during backwashing comes from elevated storage then the amount of water taken out of elevated storage would be between 112,500 and 143,750 gallons. The standpipe at the Terrace Plant holds 173,000 gallons and the elevated storage tank at the Brant Beach Plant holds 150,000 gallons; thus, the drawdown during backwashing would be rather large. However, the two plants pump water into the same distribution system and the drawdown at one plant could be limited by additional water pumped from the other plant. The most likely occurrence during backwashing would be a drawdown of water from elevated storage at the plant where the filters are being backwashed and a reversal of flow in the north-south mains in the area of the distribution system between the two plants. The reversal of flow in a large area of the distribution system is likely to be a more significant cause of increased numbers of coliforms in distribution system samples than the drawdown of water from elevated storage.

J.4.3.2 Downingtown Water Treatment Plant

The pressure in the water distribution system is completely unaffected by backwashing of filters, so this hypothesis could not be tested on this system.

J.4.3.3 Brooklawn Water Treatment Plant

The only reasonable test of the hypothesis that elevated storage tanks

TABLE J6

RESULTS OF BACTERIOLOGICAL TESTS ON FILTERED WATER SAMPLES
BROOKLAWN, NEW JERSEY

Date	Before Backwash					After Backwash				
	No.of Samp.	Coliforms (per 100ml)		Std.Plate Count (per ml)		No.of Samp.	Coliforms (per 100ml)		Std.Plate Count (per ml)	
		Mean	Std. Dev.	Mean	Std. Dev.		Mean	Std. Dev.	Mean	Std. Dev.
3/4/80	10	0	-	0.10	0.211	10	0	-	0.10	0.211
3/11/80	15	0	-	0.133	0.352	15	0	-	0.20	0.414
3/18/80	15	0	-	0.80	2.83	15	0	-	1.40	1.682
3/25/80	7	0	-	0.857	1.21	8	0	-	0.125	0.353
Overall	47	0	-	0.404	1.66	48	0	-	0.563	1.123

could be a source of coliform bacteria for the distribution system was made in Brooklawn after the sampling of the filters and backwash was completed. The separation in time was necessary to avoid overloading our laboratory capacity to process samples.

Samples from the top of the elevated storage tank were obtained by having the operator manually override the cut off on the high lift pump and allow the pump to operate until the tank overflowed. The overflow pipe was flushed out for 20 minutes before the samples were taken. The samples from the bottom of the tank were taken from the line which connects the tank to the distribution system.

The bacteriological results of the samples from the elevated storage tank are presented in Table J7. Coliforms were found only on one sampling day on samples from the top of the tank. The standard plate counts were always higher on the samples from the top of the tank than on the samples from the bottom of the tank but not significantly different.

No free chlorine residual was found in any of these samples but the total chlorine residual ranged from a trace to 0.7 mg/l. The lowest chlorine residuals were found on 4/15.

As shown in Table J8 eight of the nine coliforms from the elevated storage tank were identified as *Escherichia coli*. According to the results of the API test all of these *E. coli* isolates were the same biotype. An *E. coli* of the same biotype was found in a distribution system sample collected one block from the elevated storage tank on an earlier sampling date. This was the only *E. coli* isolate from the distribution system during the entire sampling period.

The most likely source of the coliforms found in the top of the storage tank is birds. The predominance of *E. coli* is suggestive of recent fecal contamination. It would be expected that *E. coli* would die-away rapidly in the presence of a free chlorine residual.

J.5 SUMMARY AND FINDINGS

Sampling programs carried out at four different water treatment plants provide some information on the effect filter backwashing may have on the microbiological quality of drinking water. Since the results were largely negative and the two treatment plants operated by the Long Beach Water Company are similar to the treatment plant of the Brooklawn Water Department, these findings must be considered to be tentative until or unless other treatment plants with different raw water quality and operating conditions are examined.

1. Backwash water may contain coliforms and other bacteria removed by filtration, but the coliform density and standard plate count can be expected to be very low in most cases.

TABLE J7

RESULTS OF BACTERIOLOGICAL TESTS ON SAMPLES FROM ELEVATED STORAGE TANK BROOKLAWN, NEW JERSEY

Date	Samples from Top of Tank					Samples from Bottom of Tank				
	No. of Samples	Coliforms (per 100ml)		Std. Plate Count (per ml)		No. of Samples	Coliforms (per 100ml)		Std. Plate Count (per ml)	
		Mean	Std. Dev.	Mean	Std.Dev.		Mean	Std. Dev.	Mean	Std.Dev.
4/1/80	10	0	-	1.6	1.84	10	0	-	1.1	1.19
4/8/80	10	0	-	3.3	3.16	10	0	-	1.6	1.26
4/15/80	10	0.9	0.87	3.7	2.83	10	0	-	0	-
Overall	30	0.30	0.65	2.87	2.74	30	0	-	0.9	1.18

TABLE J8

IDENTIFICATION OF COLIFORMS FROM STORAGE TANK SAMPLES

BROOKLAWN, NEW JERSEY

Date	Sample No.	Coliform Count	Colonies Isolated	Verification		IMViC	API Identification
				LTB	BGLB		
4/15	175T-2	1	1	+	+	++--	*Escherichia coli*
	175T-3	2	2	+	+	++--	*Escherichia coli*(2)
	175T-4	1	1	+	+	++--	*Escherichia coli*
	175T-5	1	1	+	+	++--	*Escherichia coli*
	175T-6	2	2	+	+	-+-+	*Citrobacter freundii*
				+	+	++--	*Escherichia coli*
	175T-7	2	2	+	+	++--	*Escherichia coli*(2)

a. When coliforms are not found in the raw water and treated water before filtration, they are not found in the backwash water (Brooklawn and the two Long Beach plants).

b. Standard plate counts on the backwash water are not significantly higher than standard plate counts on the treated water before filtration.

c. The high concentrations of suspended solids in the backwash water make the interpretation the results of fermentation tube tests questionable at best. Even when gas production in lauryl tryptose broth is observed it is difficult either to verify the presence of coliforms using brilliant green bile broth or to isolate coliforms from the original tubes.

2. A coliform was found only once in the first flush of water from a filter following backwash.

a. Clearly coliforms will not be introduced into the filter by backwashing if there are no coliforms in the water used for backwashing. This apparently was the case at Brooklawn and the two Long Beach Plants.

b. If the backwash water is accumulated in an elevated storage tank, the storage tank may be a source of coliforms which could be found in the filter immediately following backwashing.

3. Elevated storage tanks on the distribution system may contain higher numbers of coliform bacteria than most of the water in the distribution system. If drawdown of water in the elevated storage tank occurs during backwashing, coliform bacteria may be introduced into the distribution system.

*U.S. GOVERNMENT PRINTING OFFICE: 1982-0-361-082/316